



Ελληνικό Μεσογειακό Πανεπιστήμιο

Σχολή Μηχανικών

**Τμήμα: Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών**

Πτυχιακή Εργασία

Θέμα: Ηθική και Υπολογιστική Νοημοσύνη

ΦΥΝΤΡΙΛΑΚΗ ANNA (AM: 3236)

Επιβλέπων Καθηγητής: Ph.D. ΓΙΩΡΓΟΣ Μ. ΠΑΠΑΔΟΥΡΑΚΗΣ

ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΗΘΙΚΗ

*"Η επιστήμη ανακαλύπτει αυτά που υπάρχουν.
Η τεχνολογία μετατρέπει αυτήν τη γνώση
σε πράγματα που δεν υπήρξαν ποτέ"
Theodore Von Karman.*



Ευχαριστίες

Κατά κύριο λόγο, οφείλω να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα Καθηγητή μου Δρ. Γώργιο Μ. Παπαδουράκη, ο οποίος μου προσέφερε το ενδιαφέρον θέμα και την εμπιστοσύνη που μου έδειξε δίνοντάς μου τη δυνατότητα να εκπονήσω την πτυχιακή μου εργασία στο συγκεκριμένο επιστημονικό τομέα. Τον ευχαριστώ επίσης για τις πολύτιμες συμβουλές που μου παρείχε καθ' όλη τη διάρκεια της εργασίας, καθώς και για την απρόσκοπτη υποστήριξη και καθοδήγηση που μου παρείχε καθ' όλη τη διάρκεια των σπουδών μου.

Τέλος θα ήταν παράλειψή μου να μην ευχαριστήσω την οικογένειά μου για τα όσα έχει κάνει για μένα όλα αυτά τα χρόνια και τους καρδιακούς μου φίλους για τη στήριξη, τη συμπαράσταση και την κατανόησή τους.

Abstract

Ethics is becoming more and more important in our life. When artificial intelligence was initially introduced as a new trend, discussing about it was regarded utopic, since there was no ethical problem to be solved. Nowadays, however, the ethical dilemmas that emerge, regarding computational intelligence and the superintelligence of the future, are becoming so significant that necessitate the establishment of normative rules of ethical nature that will regulate the coexistence of people and machines. This work presents the role of ethics in developing computational intelligence. In addition, the scientists' concerns are listed and analyzed as well as suggestions for future actions are made. The aim is not only to investigate the ethical and moral dilemmas that computational systems may cause, but also to increase readers' awareness for more research.

Σύνοψη

Η ηθική γίνεται όλο και πιο σημαντική στην ζωή μας. Όταν η Τεχνητή Νοημοσύνη πρωτοπαρουσιάστηκε, σαν μία νέα τάση, φαινόταν σαν ουτοπία να μιλάμε για αυτήν και φυσικά δεν είχαμε κανένα ηθικό πρόβλημα για επίλυση. Πλέον τα ηθικά διλήμματα που προκύπτουν, όσον αφορά την υπολογιστική νοημοσύνη και τις υπερδοιάνες του μέλλοντος είναι τόσο σημαντικά, που καταστούν αναγκαία την δημιουργία ρυθμιστικών κανόνων ηθικής φύσεως για την συνύπαρξη ανθρώπων και μηχανών. Σε αυτήν την εργασία παρουσιάζεται ο ρόλος της ηθικής στην ανάπτυξη της υπολογιστικής νοημοσύνης, αναλύονται οι ανησυχίες, παρατίθενται απόψεις επιστημόνων, καθώς και προτάσεις για μελλοντικές ενέργειες. Στόχος είναι όχι μόνο η διερεύνηση των δεοντολογικών και ηθικών διλημάτων που προκύπτουν από τα υπολογιστικά συστήματα, αλλά και η ευαισθητοποίηση του αναγνωστικού κοινού για παραπάνω έρευνα.

Πίνακας περιεχομένων

ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΗΘΙΚΗ	ii
Λίστα Εικόνων.....	ix
1. ΕΙΣΑΓΩΓΗ.....	1
1.1 ΠΕΡΙΛΗΨΗ.....	2
1.2 Κίνητρο για την Διεξαγωγή της Εργασίας – Στόχοι	4
1.3 Δομή Εργασίας.....	5
ΚΕΦΑΛΑΙΟ 2.....	6
ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ	6
2.1. Αναφορά στην έννοια της νοημοσύνης	6
2.2. Ιστορική αναδρομή στις έξυπνες μηχανές.....	8
2.3. Τεχνητή νοημοσύνη (Artificial Intelligence – AI).....	9
2.4. Υπολογιστική Νοημοσύνη (Computational Intelligence - CI).....	10
2.4.1. Ανάλυση των κύριων συμπληρωματικών τεχνικών της υπολογιστικής νοημοσύνης.....	12
2.4.2. Απόψεις επιστημόνων για την υπολογιστική νοημοσύνη.....	14
2.5. Αποσαφήνιση των όρων Υπολογιστική Νοημοσύνη και Τεχνητή Νοημοσύνη.....	17
ΚΕΦΑΛΑΙΟ 3.....	20
ΗΘΙΚΗ	20
3.1. Εισαγωγή στην έννοια της ηθικής.	20
3.2. Ιστορική αναδρομή στην έννοια της ηθικής.	21
3.3. Ηθική της Νοημοσύνης.....	22
3.4. Ηθική της Τεχνητής Νοημοσύνης.....	23
3.4.1. Έννοια της ρομποτικής.	24
3.4.2. Ηθική της μηχανής (machine ethics).....	25
3.5. Ανάλυση της έννοιας: υπολογιστική ηθική (CE:computer ethics).	27
ΚΕΦΑΛΑΙΟ 4.....	29
ΣΥΝΔΕΣΗ ΗΘΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ	29
4.1. Η ηθική υπόσταση των υπολογιστικών συστημάτων	29
4.2 Συνείδηση και Υπολογιστική Νοημοσύνη.....	31

4.3	Αυτόνομα συστήματα και ηθική.....	33
4.4	Ηθική και παιδαγωγική χρήση των υπολογιστών.....	38
4.5	Αναγκαιότητα πρόληψης ηθικών προβλημάτων και συμπεριφορών.	40
ΚΕΦΑΛΑΙΟ 5.....		42
ΗΘΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΣΤΙΣ ΣΥΝΕΠΕΙΕΣ ΤΗΣ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ		42
5.1	Πώς αντιλαμβάνεται την εξέλιξη των υπολογιστικών συστημάτων ο άνθρωπος.....	42
5.2	Συνέπειες της εξέλιξης στην καθημερινότητα – Κοινωνικές συνέπειες.	46
5.3.	Επαγγελματικές Συνέπειες.....	52
5.4	Θεολογικές επιπτώσεις από την εξέλιξη της υπολογιστικής νοημοσύνης.	58
5.5	Επιπτώσεις στις σχέσεις των δύο φύλων.....	61
5.6	Οι συνέπειες της εξέλιξης των ιατρικών υπολογιστικών συστημάτων (θεραπευτική ιατρική χρήση ρομπότ.....	64
5.7	Συνέπειες από την εξέλιξη των στρατιωτικών υπολογιστικών συστημάτων.....	67
5.8	SUPERINTELLIGENT και ΜΟΝΑΔΙΚΟΤΗΤΑ: Μία ηθική και τεχνολογική πρόκληση.	70
5.8.1.	Ανάλυση του όρου superintelligence -υπερδιάνοια- και οι προκλήσεις που θα προκύψουν με τον ερχομό τους.	71
5.8.2.	Ανάλυση του όρου technological singularity -τεχνολογική μοναδικότητα- και οι φοβίες των ερευνητών.	73
ΚΕΦΑΛΑΙΟ 6.....		75
ΗΘΙΚΑ ΔΙΛΗΜΜΑΤΑ ΠΟΥ ΠΡΟΚΥΠΤΟΥΝ ΑΠΟ ΤΗΝ ΕΞΕΛΙΞΗ ΤΗΣ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ		77
6.1	ΕΙΣΑΓΩΓΗ.....	75
6.2	Ηθική ανάθεση στις μηχανές - Το συναίσθημα ως παράγοντας λήψης ορθής απόφασης. .	76
6.3	Ηθικά διλήμματα που προκύπτουν από την χρήση υπολογιστικών συστημάτων.	80
6.4	SUPERINTELLIGENCE και ΜΟΝΑΔΙΚΟΤΗΤΑ - Ηθικά διλήμματα που προκύπτουν.	85
6.5	Ηθικά ζητήματα που εγείρονται από τη χρήση ιατρικών υπολογιστικών συστημάτων.	90
6.6	Ηθικά ζητήματα που ανακύπτουν από την στρατιωτική χρήση των ευφυϊών συστημάτων.	94
ΚΕΦΑΛΑΙΟ 7:		98
ΔΙΚΑΙΩΜΑΤΑ ΚΑΙ ΝΟΜΙΚΗ ΕΥΘΥΝΗ ΤΩΝ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ		100
7.1	Χωρήγηση ανθρωπίνων δικαιωμάτων στα ρομπότ και ηθική.....	98
7.2	Η απόδοση νομικής ευθύνης ως ηθική αναγκαιότητα.	101

7.3 Παραχώρηση νομικής υπόστασης στα υπολογιστικά συστήματα.	104
7.4 Κινήσεις - Μέτρα που πάρθηκαν για την προστασία ανθρώπων και ρομπότ.	108
7.5 Προτάσεις επιστημόνων για μελλοντικές ενέργειες.....	116
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	121
ΒΙΒΛΙΟΓΡΑΦΙΑ:.....	126

Λίστα Εικόνων

Εικόνα 1:	11
Αλληλεπίδραση ανθρώπου - υπολογιστή σε 0-1	11
Εικόνα 2:	14
The ABCs: Neural Networks, Pattern Recognition and Intelligence (Bezdek, 1992).....	14
Εικόνα 3:	16
Simple reflex agent, based on Artificial Intelligence: A Modern Approach.....	16
Εικόνα 4:	35
Τρόμος: Ρομπότ με ιδιαίτερα υψηλή νοημοσύνη δραπέτευσε για δεύτερη φορά από εργαστήριο	35
Εικόνα 5:	38
ΝΑΟ: το κορυφαίο στον κόσμο και το πιο ευρέως χρησιμοποιούμενο ανθρωποειδές ρομπότ	38
Εικόνα 6:	40
Ρομπότ έκανε ανήθικες προτάσεις σε 7χρονη.....	40
Εικόνα 7:	47
Το εργατικό δυναμικό αντικαθίσταται από ευφυής τεχνητές νοημοσύνες.	47
Εικόνα 8:	50
Ερευνητές στην Ιαπωνία δείχνουν τον τρόπο αποκοδικοποίησης της σκέψης	50
Εικόνα 9:	51
Απεικόνιση της σκέψης των εθελοντών, από Τεχνητή Νοημοσύνη.	51
Εικόνα 10:	55
IFR 2015, Ετήσιος εφοδιασμός βιομηχανικών ρομπότ σε Ασία/Αυστραλία, Ευρώπη και Αμερική....	55
Εικόνα 11:	60
Ο ιερέας-ρομπότ «BlessU-2» ευλογεί.....	60
Εικόνα 12:	62
Ο Ρίκι Μα μαζί με την σύντροφό του – ανθρωπόμορφο ρομπότ Mark 1.	62
Εικόνα 13:	63
Ο Sergi Santos μαζί με την σύντροφό του και το ρομπότ του σεξ, Σαμάνθα	63
Εικόνα 14:	69
Ο σμήναρχος εν αποστρατεία Τζιν Λη στον εξωμοιωτή πτήσης ALPHA.....	69
Εικόνα 15:	77

Πιστεύετε ότι μπορούμε να σχεδιάσουμε ηθικά ρομπότ, ακόμα κι εάν δεν μπορούμε να συμφωνήσουμε στο τί κάνει έναν άνθρωπο ηθικό;.....	77
Εικόνα 16:	79
Το πρόβλημα του τρόλεϋ.	79
Εικόνα 17:	82
Το ανθρωποειδές ρομπότ Atlas, της Boston Dynamics, ένα από τα πιο εξελιγμένα ρομπότ που κινείται σαν άνθρωπος.....	82
Εικόνα 18:	84
Σοφία - το ρομπότ με υπηκοότητα από τη Σαουδική Αραβία.....	84
Εικόνα 19:	92
Η τεχνολογική μοναδικότητα (singularity)	92
Εικόνα 20:	94
Ποιός ευθύνεται για τα θύματα των ρομποτικών επιθέσεων.	94
Εικόνα 21:	97
Επίθεση από τηλεχειριζόμενα ρομποτικά, μη επανδρωμένα οχήματα στο Βαζιριστάν.....	97
Εικόνα 22:	115
Αυτοκινούμενο όχημα.....	115
Εικόνα 23:	120
Διανύουμε την εποχή που οι άνθρωποι πρέπει να εξετάσουν προσεκτικά πολλαπλές λεπτομέρειες πριν μπορέσουν να κρίνουν αν ένα άτομο ή ένα αντικείμενο είναι ζωντανό.....	120

1. ΕΙΣΑΓΩΓΗ

Σκοπός αυτής της πτυχιακής εργασίας είναι η ανάλυση του όρου υπολογιστική νοημοσύνη, όπως και του όρου ηθική και της σχέσης του με την υπολογιστική νοημοσύνη. Επίσης σε αυτήν την εργασία, αναφέρονται θεωρίες γνωστών επιστημόνων πάνω σε αυτούς τους όρους και παρουσιάζεται η εξέλιξη των έξυπνων μηχανών από την αρχαιότητα μέχρι σήμερα, σε τομείς όπως η ρομποτική, η τεχνητή νοημοσύνη, η ιατρική, η χρήση τους στο στρατό, στην αστυνομία και γενικά στην καθημερινότητα. Αναλύονται οι ηθικές επιπτώσεις αυτής της αλματώδους εξέλιξης στην κοινωνία, στην εργασία, στις σχέσεις των δύο φύλλων, χωρίς να παραλείπεται και μια θεολογική προσέγγιση, στο εν λόγω θέμα.

Ακόμα, επιχειρείται η καταγραφή ηθικών διλημάτων που δημιουργούνται σχετικά με τα δικαιώματα των ανθρώπων, αλλά και των υπολογιστικών συστημάτων, όπως και η αναφορά σχετικών εκτιμήσεων. Δίνονται παραδείγματα συμπεριφοράς, θετικά και αρνητικά, ανθρώπων σε ρομπότ, όπως και πράξεων έξυπνων υπολογιστικών συστημάτων και των επιπτώσεών τους στην ανθρώπινη καθημερινότητα, είτε ειρηνική είτε εμπόλεμη.

Αναφέρονται ηθικά διλήματα που προκύπτουν από τη χρήση "έξυπνων μηχανών", όπως για παράδειγμα αν έχουν ηθική και νομική υπόσταση τα superintelligent συστήματα. Επίσης, κινήσεις και μέτρα που πάρθηκαν τόσο σε ευρωπαϊκό όσο και σε παγκόσμιο επίπεδο, για την προστασία των ανθρώπων από τα ρομπότ, αλλά και το αντίθετο. Τέλος, επιχειρήθηκε να δοθεί ένα πιθανό πλάνο δράσεων για το μέλλον. Θα ήθελα επίσης να αναφέρω, ότι επειδή υπάρχει μία σύγχυση όσον αφορά τους όρους τεχνητή νοημοσύνη και υπολογιστική νοημοσύνη, προσπάθησα να αποσαφηνίσω αυτούς τους όρους, όπως επίσης, να χρησιμοποιήσω και τους δύο στην πορεία της εργασίας, αφού συχνά χρησιμοποιούνται με ταυτόσημη έννοια.

Η πτυχιακή αυτή εργασία, έρχεται να φωτίσει κάποια από τα ηθικά ζητήματα που προκύπτουν από τη χρήση της υπολογιστικής νοημοσύνης σε όλες τις εκφάνσεις της καθημερινότητας. Επειδή η υπάρχουσα γνώση και βιβλιογραφία γύρω από αυτά τα ζητήματα είναι ελάχιστη, χρησιμοποιήθηκαν πηγές, όχι μόνο από βιβλία αλλά και από εφημερίδες και ηλεκτρονικές διευθύνσεις. Οι ηθικές ανησυχίες είναι καινούριες και συνεχώς πολλαπλασιάζονται, για αυτό μέχρι σήμερα δεν μπορούν να απαντηθούν επαρκώς.

Στην προσπάθειά μου να δώσω μια πιο ρεαλιστική εικόνα της ανθρώπινης κοινής γνώμης, ως προς τα ηθικά διλήματα που προκύπτουν από την ύπαρξη έξυπνων υπολογιστικών συστημάτων, προέβησα στη σύνταξη ερωτηματολογίου στην ελληνική και αγγλική γλώσσα. Διανεμήθηκε σε διαφορετικές περιοχές της Ελλάδας, όπως και σε διάφορες χώρες του εξωτερικού, σε άτομα διαφορετικών ηλικιών, μόρφωσης και κοινωνικής θέσης, με σκοπό να έχει ένα πιο αντιπροσωπευτική εικόνα της ανθρώπινης γνώμης.

Θεωρώ ότι η συγκεκριμένη εργασία μπορεί να αποτελέσει τη βάση για μια πιο εξειδικευμένη έρευνα.

1.1 ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια η εξέλιξη της τεχνολογίας επιδρά στην ανθρώπινη φύση τόσο πολύ, που ακουμπάει τα όρια έργων επιστημονικής φαντασίας. Σε ολόκληρο τον κόσμο, απλοί άνθρωποι αλλά και επιστήμονες, αναρωτιούνται για το πώς οι υπολογιστές και οι "έξυπνες μηχανές" θα επηρεάσουν τον τρόπο ζωής μας, τις ικανότητές μας, τις γνώσεις μας, αλλά και με ποιό τρόπο θα μπορέσουν να συνεργαστούν με το ανθρώπινο μυαλό, με τον καλύτερο δυνατό τρόπο, για το καλύτερο δυνατό αποτέλεσμα. Οι ερευνητές, μας προετοιμάζουν γι' αυτή την καινούρια κατάσταση, ενώ φιλόσοφοι, πολιτικοί, και θρησκευτικοί ηγέτες, θέτουν νέα ερωτήματα, εκφράζουν φόβους και υπενθυμίζουν τα ηθικά διλήμματα που προκύπτουν από αυτή την καινούργια σχέση.

Καθώς η επιστήμη εξελίσσεται, και το διαδίκτυο διεισδύει βαθύτερα στη ζωή μας, οι ανησυχίες σχετικά με την υπολογιστική νοημοσύνη γίνονται εντονότερες. Αναπτύσσονται ολοένα και περισσότερο έξυπνα αντικείμενα, που η ευρύτερη εγκατάστασή τους δημιουργεί ερωτήματα για πιθανές βαθιές ηθικές, ψυχολογικές, κοινωνικές, οικονομικές και νομικές συνέπειες για την ανθρώπινη κοινωνία και τον πλανήτη μας.

Αυτό που είναι σίγουρο είναι ότι οι πρόσφατες εξελίξεις στην πληροφορική, μας κατευθύνουν σε ένα μοναδικό σημείο καμπής στην ανθρώπινη ιστορία. Πιθανότατα σύντομα να ανατεθεί η διαχείριση του περιβάλλοντός μας, η οικονομία, η ασφάλεια, οι υποδομές, η παραγωγή τροφίμων, η υγειονομική περίθαλψη, και σε μεγάλο βαθμό ακόμη και οι προσωπικές μας δραστηριότητες, σε τεχνητά ευφυή συστήματα υπολογιστών.

Ήδη, εξελιγμένες συσκευές σκέψης αναπτύσσονται και τοποθετούνται μέσα σε μεταλλικά αντικείμενα, τα γνωστά μας ρομπότ, τα οποία έχουν τη δυνατότητα κίνησης και ομιλίας. Σήμερα αντιμετωπίζουμε τα ρομπότ ως άβυσχα αντικείμενα και αυτό συμβαίνει επειδή ένα ρομπότ δεν διαθέτει τα απαραίτητα χαρακτηριστικά για να θεωρηθεί ζωντανό. Έχει περιορισμένη κινητικότητα, γι' αυτό και πρέπει να προγραμματιστεί από τον άνθρωπο, έλλειψη "σκέψης", αισθήσεων, συναισθημάτων, και το πιο σημαντικό δεν μπορεί να βιώσει πόνο ή φόβο. Ωστόσο, το ρομπότ του αύριο, θα έχει αναμφίβολα πολλά από αυτά τα χαρακτηριστικά και ίσως να γίνει ένας οικείος σύντροφος για τον άνθρωπο. Όταν συμβεί αυτό, θα πρέπει να έχει άραγε και τα ίδια δικαιώματα με τον άνθρωπο;

Αυτές οι συσκευές σκέψης, τα προγράμματα δηλαδή των ηλεκτρονικών υπολογιστών, ακόμα και της "τεχνητής νοημοσύνης", κατασκευάστηκαν με σκοπό να ελέγχονται από τον άνθρωπο. Μερικοί επιστήμονες υποστηρίζουν ότι αν οι υπολογιστές κατασκευαστούν έτσι ώστε να ενεργούν από μόνοι τους, θα μπορούν να επιταχύνουν και να μεγεθύνουν τα λάθη μας και να κάνουν μεγαλύτερη ζημιά από ό,τι ένα άτομο είναι ικανό να κάνει. Αυτό γιατί σκέπτονται και ενεργούν πολύ πιο γρήγορα από τον άνθρωπο και χωρίς να έχουν "δεύτερη σκέψη". Αυτή η σκέψη οδηγεί κάποιους επιστήμονες να πιστέψουν ότι η υπολογιστική νοημοσύνη είναι εύκολο να θέσει σε κίνδυνο την κοινωνία μας. Είναι αλήθεια ότι στο υποσυνείδητο όλων μας, χρόνια τώρα, υπάρχει ο φόβος ότι τα ρομπότ μπορούν να γίνουν εντελώς αυτόνομα, με ελεύθερη βούληση, ευφυΐα και συνείδηση. Ακόμη και ότι μπορούν να επαναστατήσουν εναντίον μας, σαν τέρατα. Όπως στις ταινίες "The Matrix", "Η Οδύσσεια του Διαστήματος", "Ο εξολοθρευτής" και "Εγώ, το ρομπότ". Κοινός παρονομαστής είναι ένας κόσμος που την κυριαρχία την έχουν οι μηχανές, κι όχι οι άνθρωποι. Για την ακρίβεια, οι "έξυπνες" μηχανές, συνήθως, στρέφονται εναντίον του δημιουργού τους. Αρκετοί επιστήμονες υποστηρίζουν ότι οι λόγοι που δεν μπορούμε ακόμη να ισχυριστούμε ότι τα ρομπότ μπορούν να κάνουν το σωστό, είναι ότι δεν είναι πλήρως αξιόπιστα και αποτελεσματικά, δεδομένου του τρόπου που κατασκευάζονται τώρα. Αναρωτιόμαστε λοιπόν, αν αλλάξει ο τρόπος προγραμματισμού τους, θα μπορούν στο μέλλον να κάνουν το σωστό; Κι αν ναι, θα το κάνουν;

Την 5η Οκτωβρίου 1994, ο Vernor Vinge (καθηγητής στο San Diego State University) έδωσε διάλεξη στο MIT με τίτλο: "The Coming Technological Singularity: How to Survive in the Post-Human Era" (Η επερχόμενη Τεχνολογική Μοναδικότητα: Πώς να επιζήσεις στην μετά-ανθρώπινη Εποχή). Περιληπτικά είπε: *"Μέσα σε τριάντα χρόνια, θα έχουμε τα τεχνολογικά μέσα για να δημιουργήσουμε υπεράνθρωπη ευφυΐα. Λίγο μετά, ο άνθρωπος της εποχής θα σταματήσει να υπάρχει. Μπορεί να αποφευχθεί η Μοναδικότητα; Αν όχι, μπορούν να καθοδηγηθούν τα γεγονότα, έτσι ώστε να μπορέσουμε να επιβιώσουμε; Τι σημαίνει επιβίωση ακόμη και σε μια μετά-ανθρώπινη Εποχή;"*¹

Τι γίνεται λοιπόν με τη δεοντολογία που αφορά τη σχέση ρομπότ-ανθρώπου; Μήπως χρειαζόμαστε ηθικούς κώδικες, για εμάς και για τα ρομπότ; Οι ρομποτικές τεχνολογίες έχουν πλέον ωριμάσει και ετοιμάζονται να φύγουν από τα ερευνητικά εργαστήρια και να έρθουν στην αγορά σε μεγάλους αριθμούς. Τα συστήματα αυτά έχουν τη δυνατότητα να φέρουν επανάσταση στην καθημερινή μας ζωή και να μεταμορφώσουν τον κόσμο μας, με γοργούς ρυθμούς. Πρέπει να είμαστε έτοιμοι γι' αυτά τα συστήματα. Πρέπει οι νομοθέτες να μπορούν να δημιουργήσουν καλύτερους κανόνες για αυτά, και οι μηχανικοί να τα σχεδιάσουν με τέτοιο τρόπο που να προστατεύονται οι αξίες της κοινωνίας μας.

Εκ των πραγμάτων, δημιουργούνται σοβαρά ερωτήματα που πρέπει να απαντηθούν. Για παράδειγμα θα έχει ένα "έξυπνο" ρομπότ το δικαίωμα να μεγαλώσει ένα παιδί; Το δικαίωμα ψήφου; Βασικά δικαιώματα όπως την προστασία της ιδιωτικής ζωής, την ασφάλεια, την ελευθερία, κ.λπ.; Τα αποτελέσματα των λαθεμένων πράξεων ενός ρομπότ, θα επιβαρύνουν αυτό ή τον προγραμματιστή του; Εάν ένα ρομπότ δημιουργήσει, κατόπιν εντολής, ένα αριστούργημα, τα πνευματικά δικαιώματα θα είναι δικά του, ή του ιδιοκτήτη του; Ποια θα είναι η διαχωριστική γραμμή μεταξύ ανθρώπου και μηχανής; Άραγε τι θα συμβεί στην κοινωνία μας, όταν τα ρομπότ θα είναι μέρος της ζωής μας;

Η προοπτική για την ανάπτυξη των ολοένα και πιο αυτόνομων συστημάτων εγείρει πολλά και πολύπλοκα ανησυχητικά ερωτήματα. Πώς θα ανταποκριθεί η κοινωνία όταν ρομπότ εκτοπίσουν τους εργαζόμενους και τις υπαλληλικές θέσεις; Τα οφέλη αυτής της τεχνολογικής επανάστασης θα διανεμηθούν ή θα συγκεντρωθούν σε λίγους τυχερούς; Πώς μπορούμε να διασφαλίσουμε ότι τα συστήματα αυτά θα σέβονται τις ηθικές αρχές μας, όταν λαμβάνουν αποφάσεις σε ταχύτητες και λογικές που υπερβαίνουν την ικανότητά μας να κατανοήσουμε; Θα χρειαστεί να τους χορηγήσουμε νομικά δικαιώματα; Τελικά, τα προϊόντα υπολογιστικής νοημοσύνης θα πρέπει να τα θεωρούμε ως εξελιγμένα άψυχα εργαλεία ή ως μία νεοεμφανιζόμενη μορφή ζωής;

¹ ©1993 by Vernor Vinge (<https://adsabs.harvard.edu/abs/1993vise.nasa...11V>)

1.2 Κίνητρο για την Διεξαγωγή της Εργασίας – Στόχοι

Αυτό που με οδήγησε να επιλέξω το συγκεκριμένο θέμα ήταν η γοητεία των γοργά εξελισσόμενων υπολογιστικών συστημάτων σε συσχετισμό με την ανθρώπινη καθημερινότητα και τους ηθικούς φραγμούς, υποχρεώσεις και δικαιώματα, που είναι ανάγκη να επιβληθούν και να προκαθοριστούν, ώστε η συμβίωση έξυπνων μηχανών και ανθρώπων, να γίνει η καλύτερη δυνατή.

Στόχος μου είναι η έρευνα των ηθικών διλημάτων όσο αφορά την υπολογιστική νοημοσύνη και πώς θα μπορούσαν να επιτευχθεί η συμβίωση αυτών των ταχύτατα εξελισσόμενων μηχανών με τους ανθρώπους. Τί ενέργειες έχουν παρθεί, είναι προγραμματισμένες να παρθούν και τί παραπάνω ενέργειες θα έπρεπε να πραγματοποιηθούν. Επίσης μέσω αυτής της πτυχιακής θέλω να ενημερώσω και να δώσω κίνητρο, ακόμα και σε άτομα που δεν έχουν επαφή με αυτόν τον τομέα, για τους ηθικούς κινδύνους που έχουν ήδη προκύψει, σκεπτόμενοι τα ηθικά διλήματα που ήδη έχουν αρχίσει να υπάρχουν πάνω σε αυτόν τον τομέα.

1.3 Δομή Εργασίας

Στην Δεύτερη ενότητα της εργασίας, γίνεται αναφορά στην υπολογιστική νοημοσύνη. Αρχικά γίνεται μία εισαγωγή στην έννοια της Νοημοσύνης γενικά και στη συνέχεια στους όρους Τεχνητή και Υπολογιστική Νοημοσύνη. Αφού γίνεται μια ιστορική αναδρομή στις έξυπνες μηχανές, από την αρχαιότητα μέχρι σήμερα, μία αναφορά στις κύριες συμπληρωματικές τεχνικές της υπολογιστικής νοημοσύνης, όπως και στις απόψεις επιστημόνων για την υπολογιστική νοημοσύνη, επιχειρείται μια αποσαφήνιση των όρων Υπολογιστική και Τεχνητή Νοημοσύνη, επειδή πολύ συχνά αυτοί οι δύο όροι χρησιμοποιούνται εκφράζοντας ακριβώς την ίδια έννοια.

Η Τρίτη ενότητα πραγματεύεται την έννοια της Ηθικής. Ύστερα από μια ιστορική αναδρομή σε αυτήν, αναλύονται οι όροι Ηθική της Νοημοσύνης, Ηθική της Τεχνητής Νοημοσύνης, συμπεριλαμβανομένων της Ρομποτικής και της Ηθικής της Μηχανής, καταλήγοντας στην ανάλυση της έννοιας της Υπολογιστικής Ηθικής.

Στην Τέταρτη ενότητα επιχειρείται η σύνδεση της Ηθικής, με την Υπολογιστική Νοημοσύνη. Εστιάζεται στην ηθική υπόσταση των υπολογιστικών συστημάτων, στη σύνδεση της ηθικής με τα αυτόνομα συστήματα και με την παιδαγωγική χρήση των υπολογιστών, αλλά και στην αναγκαιότητα της πρόληψης πιθανής δημιουργίας ηθικών προβλημάτων και συμπεριφορών.

Στην Πέμπτη ενότητα πραγματοποιείται μια ηθική προσέγγιση στις συνέπειες της εξέλιξης της υπολογιστικής νοημοσύνης. Αρχικά σχολιάζεται το πώς αντιλαμβάνεται ο άνθρωπος την εξέλιξη των υπολογιστικών συστημάτων. Κατόπιν, αναλύονται οι συνέπειες αυτής της εξέλιξης, σε κοινωνικό, επαγγελματικό, θεολογικό, ιατρικό, στρατιωτικό, ακόμα και ερωτικό επίπεδο. Τέλος, παρατίθενται οι σημαντικοί ορισμοί, Superintelligence (υπερδιάνοια) και Technological Singularity (τεχνολογική μοναδικότητα), και αναλύονται οι προκλήσεις που θα προκύψουν από τον ερχομό τους, όπως και οι φοβίες των ερευνητών.

Η Έκτη ενότητα εστιάζει στα ηθικά διλήμματα που προκύπτουν από αυτήν την εξέλιξη, σε όλα τα προαναφερθέντα στην προηγούμενη ενότητα, επίπεδα της ζωής του ανθρώπου. Αναλυτικότερα, γίνεται μια αρχική εισαγωγή του αναγνώστη σε γενικότερα ηθικά διλήμματα που προκύπτουν και κατόπιν επιχειρείται να γίνει μια ολοκληρωμένη προσέγγιση όσον αφορά τη δυνατότητα ηθικής ανάθεσης στις μηχανές, εξετάζοντας αν το συναίσθημα που δε διαθέτουν, είναι ένας σημαντικός παράγοντας για τη λήψη ορθών αποφάσεων. Στη συνέχεια τα ηθικά διλήμματα εστιάζονται σε επίπεδο χρήσης υπολογιστικών συστημάτων, σε επίπεδο Superintelligence και τεχνολογικής μοναδικότητας, ιατρικής και στρατιωτικής χρήσης τους.

Η Έβδομη ενότητα πραγματεύεται το ερώτημα αν τα υπολογιστικά συστήματα πρέπει να έχουν δικαιώματα, αλλά και νομική ευθύνη, όπως και την αναγκαιότητα δημιουργίας ηθικών ρυθμιστικών κανόνων. Επίσης, εξετάζεται πόσο ηθικό είναι να δοθούν ανθρώπινα δικαιώματα και νομική υπόσταση στα ρομπότ. Τέλος, παρατίθενται κινήσεις και μέτρα που πάρθηκαν για την προστασία των ρομπότ αλλά και των ανθρώπων, όπως και προτάσεις επιστημόνων, για μελλοντικές ενέργειες στον τομέα αυτό.

Η εργασία ολοκληρώνεται με την Όγδοη ενότητα, η οποία αναφέρεται περιληπτικά στα συμπεράσματα που προκύπτουν από αυτή την εργασία.

ΚΕΦΑΛΑΙΟ 2

ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ

2.1. Αναφορά στην έννοια της νοημοσύνης

Ο Γάλλος ψυχολόγος A. Binet, που μαζί με τον Theodore Simon το 1905 κατασκεύασαν την πρώτη κλίμακα για τη μέτρηση νοημοσύνης στα παιδιά, υποστηρίζει ότι η νοημοσύνη είναι η γενική ικανότητα του ατόμου να προσαρμόζεται σε νέες καταστάσεις. Η καλή κρίση, η καλή κατάσταση, η λογική σκέψη αποτελούν τα κύρια συστατικά στοιχεία της νοημοσύνης.

Οι ίδιοι, εισήγαγαν την ιδέα της νοητικής ηλικίας, προσδιορίζοντάς την από το μέσο όρο της χρονολογικής ηλικίας και της νοητικής ικανότητας. Ο Αμερικανός ψυχολόγος David Wechsler, ένας από τους γνωστότερους δημιουργούς δοκιμασιών νοημοσύνης που ανέπτυξε γνωστές κλίμακες νοημοσύνης, όπως η Wechsler Adult Scale Intelligence (WAIS) και η Wechsler Intelligence Scale for Children (WISC), πιστεύει ότι νοημοσύνη είναι η γενική ικανότητα του ατόμου να δρα σκόπιμα, να σκέφτεται λογικά και να αντιμετωπίζει με αποτελεσματικότητα το περιβάλλον του. Αργότερα η άποψη ότι η νοημοσύνη είναι μία γενική πνευματική ικανότητα αμφισβητήθηκε.

Η νοημοσύνη έχει μια πολυσύνθετη δομή. Επηρεάζεται από την ηλικία, την εμπειρία, τις πνευματικές ικανότητες του ατόμου, αλλά και από εξωγενείς παράγοντες. Αυτός που πρώτος προσπάθησε να μετρήσει την νοημοσύνη ήταν ο Άγγλος Sir Francis Galton (1822-1911). Το 1860 προσπάθησε να αποδείξει ότι η νοημοσύνη είναι κληρονομική, δημιουργώντας ένα τεστ που αποτέλεσε τη βάση για όλα τα άλλα που δημιουργήθηκαν μεταγενέστερα.

Πολλοί επιστήμονες με έρευνές τους, απέδειξαν ότι η κληρονομικότητα αποτελεί τον ισχυρότερο παράγοντα που καθορίζει την ευφυΐα σε ένα άτομο. Βέβαια, η νοημοσύνη απαιτεί το κατάλληλο περιβάλλον για να εξελιχτεί. Αν ένα άτομο θα φτάσει στα όρια των δυνατοτήτων της νοημοσύνης που του προσφέρουν τα γονίδια του, θα εξαρτηθεί από τα ερεθίσματα που θα δεχτεί από το οικογενειακό και σχολικό του περιβάλλον, τα κίνητρα και τις ευκαιρίες για μόρφωση που θα του δοθούν. Αυτά ενισχύουν την υπάρχουσα νοημοσύνη.

Το 1983 ο γερμανικής καταγωγής αμερικανός πολίτης Howard Gardner, που θεωρείται κορυφή στα θέματα της γνωστικής ψυχολογίας, δημοσίευσε το βιβλίο του "Frames of Mind: The Theory of Multiple Intelligences" ("Τα πλαίσια του νου: Η Θεωρία της Πολλαπλής Νοημοσύνης"), εκφράζοντας για πρώτη φορά μια θεωρία, η οποία άλλαξε τον χάρτη της εκπαίδευσης σε παγκόσμιο επίπεδο, αυτή των πολλαπλών ευφυϊών. Ο Γκάρντερ, καθηγητής Ψυχολογίας της Μάθησης στο Χάρβαρντ, καθηγητής Ιατρικής και Νευρολογίας στην Ιατρική Σχολή του Πανεπιστημίου της Βοστώνης, συγγραφέας 18 βιβλίων που έχουν μεταφραστεί σε 21 γλώσσες, εκατοντάδων άρθρων σχετικά με θέματα ψυχολογίας και τιμημένος με περισσότερα από 20 διεθνή βραβεία, ήταν ο άνθρωπος που απέδειξε ότι η νοημοσύνη μας είναι πολύ περισσότερα πράγματα από αυτό που ορίζει η «στεία» μέτρηση του IQ μας. Και έδωσε έτσι στους ειδικούς το «κλειδί» που ανοίγει την «πόρτα» της ολοκλήρωσης και της ισορροπίας του ατόμου.²

² H. Gardner, «Frames of Mind: The Theory of Multiple Intelligences», 1993.

Ανέπτυξε λοιπόν, τη θεωρία της πολλαπλής νοημοσύνης (multiple intelligence). Σύμφωνα με αυτή, όλοι ανήκουμε σε κάποια από τις παρακάτω επτά κατηγορίες νοημοσύνης:

- Τη Κινησθητική Νοημοσύνη (body-kinesthetic). Οι άνθρωποι που χαρακτηρίζονται από αυτό το είδος νοημοσύνης, μαθαίνουν με τις σωματικές εκφράσεις, και τους αρέσει να δουλεύουν με τα χέρια τους, και με ό,τι άλλο έχει κίνηση, γιατί δεν μπορούν να κάτσουν ήσυχοι. Χρησιμοποιούν το ένστικτό τους και δίνουν λύσεις σε θέματα που πολλές φορές δε γνωρίζουν.
- Τη Διαπροσωπική Νοημοσύνη (interpersonal). Είναι αυτή με την οποία οι άνθρωποι έχουν δυνατή προσωπικότητα, είναι πολύ κοινωνικοί, είναι εξωστρεφείς, έχουν μία ιδιαίτερη ευαισθησία, όπως για παράδειγμα οι κοινωνικοί λειτουργοί και οι ηθοποιοί.
- Την Ενδοπροσωπική Νοημοσύνη (intrapersonal). Αυτοί που χαρακτηρίζονται από αυτό το είδος νοημοσύνης, είναι δυνατές προσωπικότητες, γνωρίζουν πολύ καλά τον εαυτό τους, έχουν πειθαρχία και τους αρέσει να δουλεύουν μόνοι τους (πχ ηγέτες και φιλόσοφοι).
- Τη Γλωσσική Νοημοσύνη (linguistic). Αυτή τη νοημοσύνη έχουν οι ποιητές, οι συγγραφείς και οι άνθρωποι που γνωρίζουν πολλές ξένες γλώσσες.
- Τη Λογικομαθηματική Νοημοσύνη (logical-mathematical). Οι άνθρωποι που κατέχουν αυτό το είδος νοημοσύνης, σκέπτονται λογικά και έχουν την ικανότητα να βλέπουν μια κίνηση ή μια εφεύρεση, πριν καν να την δημιουργήσουν. Έχουν υψηλό δείκτη νοημοσύνης, είναι καλοί μαθητές και καλοί εφευρέτες. Σε αυτή την κατηγορία ανήκουν ο Άλμπερτ Αϊνστάιν και ο Μπιλ Γκέιτς.
- Τη Μουσική Νοημοσύνη (musical). Οι άνθρωποι με μουσική νοημοσύνη, έχουν μεγάλη κλίση στη μουσική, έχουν εξαιρετικό ρυθμό και θυμούνται μελωδίες που άκουσαν 2-3 φορές μόνο. Πολλές φορές συγκεντρώνονται καλύτερα ακούγοντας μουσική. Καμιά φορά η ευφυΐα αυτή, συνδυάζεται με μαθηματική ευφυΐα.
- Τη Χωροταξική Νοημοσύνη (spatial). Όποιος ανήκει σε αυτήν την κατηγορία έχει την ικανότητα να βλέπει σε εικόνες και να σκέφτεται σε 3 διαστάσεις. Αυτή τη νοημοσύνη έχουν οι ζωγράφοι και οι γλύπτες, οι πιλότοι.
- Τη Νατουραλιστική Νοημοσύνη (naturalist). Χαρακτηρίζει όσους έχουν την ικανότητα να αναγνωρίζουν τους ζωντανούς οργανισμούς, ενώ είναι και ευαίσθητοι στα ερεθίσματα του φυσικού κόσμου (σύννεφα, βραχώδεις σχηματισμούς, λιβάδια). Η ευφυΐα αυτή ήταν ιδιαίτερα πολύτιμη κατά τη διάρκεια της ανθρώπινης εξέλιξης μας, τότε που λειτουργούσαμε ως κυνηγοί και συλλέκτες. Εξακολουθεί να παίζει σημαντικό ρόλο για ανθρώπους όπως οι βοτανολόγοι και οι μάγειροι. Από την άλλη η καταναλωτική κοινωνία μας αξιοποιεί τη νατουραλιστική ευφυΐα, την οποία ενεργοποιεί στο να διαλέξει ανάμεσα σε αυτοκίνητα, σε παπούτσια και σε καλλυντικά.
- Την Υπαρξιακή Νοημοσύνη (existential). Οι άνθρωποι με ανεπτυγμένη υπαρξιακή νοημοσύνη έχουν μία ιδιαίτερη ευαισθησία και ικανότητα για διερεύνηση βαθυστόχαστων ερωτημάτων σχετικά με την ανθρώπινη ύπαρξη, το νόημα της ζωής, γιατί πεθαίνουμε και πώς βρεθήκαμε στην ζωή.

Η θεωρία του Gardner παρείχε τη βάση για τη μεταγενέστερη θεωρία των Salovey και Mayer, γύρω από την Συναισθηματική Νοημοσύνη "Emotional Intelligence (E.I)".³

Βέβαια, υπάρχει και μία λιγότερο γνωστή κατηγορία ευφυΐας, η οποία ονομάζεται «ηθική νοημοσύνη» (Hass 1998). Σε αυτήν την κατηγορία συνδυάζεται η «δύναμη του χαρακτήρα» και η «ορθή κρίση». Ένα είδος αρετής που «μαθαίνει» στα ανθρώπινα όντα, να κάνουν ό,τι οφείλουν και όχι ότι τους είναι αρεστό. Η ικανότητα δηλαδή του να διακρίνει κανείς μεταξύ σωστού και λάθους αλλά και να σέβεται τις αξίες των άλλων.

³ Mayer, J.D.; Salovey, P.; Caruso, D.L.; Sitarenios, G. (2001). «Emotional intelligence as a standard intelligence». *Emotion* 1: 232–242. [doi:10.1037/1528-3542.1.3.232](https://doi.org/10.1037/1528-3542.1.3.232).

2.2. Ιστορική αναδρομή στις έξυπνες μηχανές.

Από την αρχαιότητα είχαν κάνει την εμφάνισή τους οι σκεπτόμενες μηχανές. Στην ελληνική μυθολογία για παράδειγμα, αναφέρονται συστήματα όπως ο Τάλως της Κρήτης, ένας μυθικός χάλκινος γίγαντας-ρομπότ που προστάτευε την μινωική Κρήτη, και είχε κατασκευαστεί από τον Ήφαιστο. Ο Ήρωνας ο Αλεξανδρεύς, πολυπράγμων μηχανικός, άριστος μαθηματικός και μοναδικός εφευρέτης, στο βιβλίο του "Πνευματικά" περιγράφει ογδόντα περίπου αυτοματισμούς, όπως τον πρώτο αυτόματο πωλητή της ιστορίας με κέρμα και το πρώτο λειτουργικό ρομπότ της ανθρωπότητας, ένα ανθρωποειδές που κερνούσε κρασί και νερό. Επίσης, στο σύγγραμμά του "Αυτοματοποιητική", μας παρουσιάζει ένα αυτοκινούμενο προγραμματιζόμενο όχημα και το εκπληκτικό αυτόματο θέατρο με κινούμενη εικόνα και ήχο. Επίσης, δημιούργησε την πρώτη μετάδοση μηνυμάτων γράμμα-γράμμα μεταξύ υψωμάτων, με ψηφιακό σύστημα προδρομικό της σημερινής τεχνολογίας.⁴

Πολύ αργότερα, το 1614, ο John Napier, Σκωτσέζος μαθηματικός, φυσικός και αστρονόμος, δημιούργησε μια μηχανή για τον υπολογισμό των λογαρίθμων. Βέβαια, η πρώτη αληθινή αριθμομηχανή κατασκευάστηκε από τον Γάλλο Blaise Pascal το 1645, που έδινε τη δυνατότητα μαθηματικών υπολογισμών και λειτουργούσε με 5,6 ή 8 γρανάκια και εκτελούσε πρόσθεση και αφαίρεση. Η εφεύρεση του ηλεκτρονικού υπολογιστή όμως, βασίστηκε στη θεωρία υπολογισμού του μαθηματικού Alan Turing που πρότεινε ότι μια μηχανή με τη χρήση συμβόλων όπως το "0" και το "1", (κλειστός διακόπτης, ανοιχτός διακόπτης), θα μπορούσε να προσομοιώσει οποιαδήποτε πράξη μαθηματικής επαγωγής.

Ο Alan Turing ήταν και ο πρώτος που περιέγραψε μια υπολογιστική μηχανή και τον τρόπο που θα μπορούσε να πραγματοποιεί κάθε υπολογισμό και να αποθηκεύει στη μνήμη του ένα πρόγραμμα, σε ένα άρθρο που δημοσιεύθηκε το 1936 στο "Computable Numbers, with an Application to the Entscheidungsproblem".⁵ Αυτό ενέπνευσε μια ομάδα ερευνητών που άρχισε να εξετάζει το ενδεχόμενο δημιουργίας ενός ηλεκτρονικού εγκεφάλου. Το 1937 σχεδιάστηκε το πρώτο μηχάνημα που μπορεί να χαρακτηριστεί σαν ηλεκτρονικός υπολογιστής και ήταν ο ABC (Atanasoff Berry Computer). Το 1946 δημιουργήθηκε ο ENIAC που χρησιμοποιούσε ηλεκτρονικές λυχνίες αντί για μηχανικά μέρη και χρησιμοποιήθηκε στο στρατό. Αποτέλεσμα αυτών των προσπαθειών ήταν να κατασκευαστούν μερικά σημαντικά υπολογιστικά συστήματα με σημαντικότερο το UNVAC 1 (UNIVersal Automatic Computer). Στη δεκαετία του 1950 έχουμε τη δεύτερη γενιά υπολογιστών, το 1964 ξεκινά η τρίτη γενιά και σήμερα έχουμε τους υπολογιστές της πέμπτης γενιάς, με πιο ισχυρό αυτόν της IBM, με το όνομα Roadrunner. Κάποιοι ερευνητές υποστηρίζουν ότι η τεχνολογία τρέχει τόσο πολύ που αναπτύσσεται με γεωμετρική πρόοδο.

Στις μέρες μας, οι μηχανές έχουν αποκτήσει νοημοσύνη. Τη μηχανική νοημοσύνη. Σκέφτονται και λειτουργούν μόνες τους, αποκτούν ανθρώπινες ικανότητες, μαθαίνουν από τις εμπειρίες τους όπως οι άνθρωποι, προσαρμόζονται εύκολα και εξελίσσονται πολύ γρήγορα. Υπάρχουν δύο τύποι μηχανικής νοημοσύνης: η τεχνητή, βασισμένη στις σκληρές υπολογιστικές τεχνικές και η υπολογιστική, βασισμένη σε μεθόδους μαλακών υπολογιστών, οι οποίες επιτρέπουν την προσαρμογή σε πολλές καταστάσεις.

⁴ Ταξίδι στην αρχαία Ελλάδα: Οι Σημαντικότερες Εφευρέσεις Των Αρχαίων Ελλήνων, *archaia-ellada.blogspot.gr*.

⁵ Εμφανίστηκε στα Πρακτικά της Μαθηματικής Εταιρείας του Λονδίνου, Σειρά 2, τόμος 42, σελ. 230 - 265, «What is o Turing Machine» by Jack Copeland - Copyright B.J. Copeland), Ιούλιος 2000.

2.3. Τεχνητή νοημοσύνη (Artificial Intelligence – AI)

Το 1955 ο John McCarthy, Επίκουρος Καθηγητής Μαθηματικών στο Dartmouth College, αποφάσισε να οργανώσει μια ομάδα για να διευκρινίσει και να αναπτύξει ιδέες σχετικά με τις μηχανές σκέψης. Για το νέο αυτό πεδίο διάλεξε το όνομα "Τεχνητή Νοημοσύνη", και αυτό εν μέρει, λόγω της ουδετερότητάς του.

Στις αρχές του 1955, ο McCarthy πλησίασε το Ίδρυμα Rockefeller για να ζητήσει χρηματοδότηση για ένα θερινό σεμινάριο στο Dartmouth, για περίπου 10 συμμετέχοντες. Τον Ιούνιο, ο ίδιος και ο Claude Shannon, ιδρυτής της Θεωρίας της Πληροφορίας στη Bell Labs, συναντήθηκαν με τον Robert Morison, Διευθυντή Βιολογικών και Ιατρικών Ερευνών, για να συζητήσουν την ιδέα και την πιθανή χρηματοδότηση, παρόλο που ο Morison δεν ήταν σίγουρος εάν θα διατεθούν χρήματα για ένα οραματικό έργο.⁶

Στις 2 Σεπτεμβρίου του 1955, το έργο προτάθηκε επίσημα από τους McCarthy, Marvin Minsky (Αμερικανός επιστήμονας που ασχολήθηκε με την έρευνα πάνω στην τεχνητή νοημοσύνη), Nathaniel Rochester (αμερικάνος επιστήμονας που ασχολήθηκε με την έρευνα της τεχνητής νοημοσύνης και εφεύρε τον πρώτο συμβολομεταφραστή) και Claude Shannon. Η πρόταση πιστώνεται με την εισαγωγή του όρου «τεχνητή νοημοσύνη» και χαρακτηριστικά αναφέρει: *"Προτείνουμε να διεξαχθεί μια μελέτη τεχνητής νοημοσύνης 2 μηνών και 10 ατόμων κατά τη διάρκεια του καλοκαιριού του 1956 στο κολέγιο Dartmouth στο Ανόβερο του Νιου Χάμσαϊρ. Η μελέτη θα προχωρήσει με βάση την εικασία ότι κάθε πτυχή της μάθησης ή οποιουδήποτε άλλου χαρακτηριστικού της νοημοσύνης, μπορεί να περιγραφεί με τόσο μεγάλη ακρίβεια ώστε να μπορεί να γίνει μια μηχανή για να την προσομοιώσει. Θα γίνει μια προσπάθεια να βρεθεί πώς θα πρέπει να κατασκευαστούν τα μηχανήματα ώστε να χρησιμοποιούν τη γλώσσα, να κάνουν πράξεις και να σχηματίζουν έννοιες, να λύνουν όλα τα είδη προβλημάτων που ένας άνθρωπος θα μπορούσε να έχει και να βελτιώνονται. Θεωρούμε ότι μπορεί να σημειωθεί σημαντική πρόοδος σε ένα ή περισσότερα από αυτά τα προβλήματα εάν μια προσεκτικά επιλεγμένη ομάδα επιστημόνων εργαστούν μαζί πάνω σε αυτό για ένα αποτέλεσμα".*⁷

Η λέξη «τεχνητή νοημοσύνη» χρησιμοποιείται επιμελώς όταν ένα μηχανήμα μιμείται τις «γνωστικές» λειτουργίες του ανθρώπινου μυαλού, όπως είναι η «μάθηση» και η «επίλυση προβλημάτων».⁸ Το πεδίο βασίστηκε στον ισχυρισμό που αναφέραμε παραπάνω, δηλαδή στο ότι η ανθρώπινη νοημοσύνη «μπορεί να περιγραφεί με τόση ακρίβεια, ώστε να μπορεί να γίνει μια μηχανή για να την προσομοιώσει». Αυτό εγείρει φιλοσοφικά επιχειρήματα σχετικά με τη φύση του νου και την ηθική της δημιουργίας τεχνητών όντων που έχουν ανθρωποειδή νοημοσύνη, θέματα που έχουν διερευνηθεί από τον μύθο, τη μυθοπλασία και τη φιλοσοφία από την αρχαιότητα.

Σύμφωνα με τους Barr και Feigenbaum, η Τεχνητή Νοημοσύνη είναι ο τομέας της επιστήμης των υπολογιστών, που ασχολείται με τη σχεδίαση ευφώνων (νοημόνων) υπολογιστικών συστημάτων, δηλαδή συστημάτων που επιδεικνύουν χαρακτηριστικά που σχετίζουμε με τη νοημοσύνη στην ανθρώπινη συμπεριφορά. Ο Douglas Hofstadter, Αμερικανός καθηγητής γνωστικής επιστήμης, προτείνει ότι η νοημοσύνη είναι, να ανταποκρίνεται σε καταστάσεις με ελαστικότητα (όχι μηχανική συμπεριφορά), να κατανοεί

⁶ Kline, Ronald R., *Cybernetics, «Automata Studies and the Dartmouth Conference on Artificial Intelligence»*, IEEE Annals of the History of Computing, Οκτώβριος-Νοέμβριος 2011, IEEE Computer Society.

⁷ McCarthy, J., Minsky, M., Rochester, N., Shannon, C.E., *«A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence»*, Αύγουστος 1955.

⁸ Russell & Norvig 2009, p.2.

τα ασαφή ή αντιφατικά μηνύματα από τα συμφραζόμενα, να αναγνωρίζεις και να ιεραρχείς τα διάφορα δεδομένα με βάση τη σπουδαιότητά τους, να βρίσκεις ομοιότητες μεταξύ καταστάσεων οι οποίες μοιάζουν διαφορετικές και να βρίσκεις διαφορές μεταξύ καταστάσεων οι οποίες μοιάζουν παρόμοιες.

Οι ικανότητες αυτές έχουν τουλάχιστον ένα κοινό χαρακτηριστικό: αποκτώνται εύκολα από τους ανθρώπους και συνήθως τις αποκαλούμε κοινή λογική. Έχει αποδειχθεί ότι όσο πιο απλή και αυτονόητη για τους ανθρώπους είναι μια λειτουργία τόσο πιο δύσκολα μπορούμε να τη μεταφέρουμε σε έναν υπολογιστή. Άλλοι ορισμοί επικεντρώνονται στη διαδικασία σκέψης και συλλογισμού και άλλοι στη συμπεριφορά. Υπάρχουν ορισμοί σύμφωνα με τους οποίους στόχος της Τεχνητής Νοημοσύνης είναι:

α) να φτιάξει συστήματα που να σκέφτονται όπως οι άνθρωποι. Η προσπάθεια δηλαδή να κατασκευάσουμε υπολογιστές με διανοητική ικανότητα με την πλήρη και κυριολεκτική έννοια του όρου.⁹

β) να συμπεριφέρονται όπως οι άνθρωποι. Με άλλα λόγια, η μελέτη του πώς να κάνουμε τους υπολογιστές να κάνουν πράγματα που αυτήν τη στιγμή οι άνθρωποι είναι καλύτεροι.¹⁰

γ) να σκέφτονται λογικά. Η μελέτη δηλαδή των υπολογισμών που καθιστούν εφικτή την αντίληψη, λογική σκέψη και αντίδραση.¹¹

δ) να αντιδρούν λογικά. Αυτό έχει σαν στόχο ο τομέας της επιστήμης των υπολογιστών που ασχολείται με την αυτοματοποίηση της ευφυούς συμπεριφοράς.¹²

Ένας γενικός ορισμός θα μπορούσε να είναι ότι η Τεχνητή Νοημοσύνη είναι ο τομέας της Επιστήμης των Υπολογιστών που ασχολείται με τη σχεδίαση και την υλοποίηση προγραμμάτων τα οποία είναι ικανά να μιμηθούν τις ανθρώπινες γνωστικές ικανότητες. Εμφανίζουν με αυτόν τον τρόπο χαρακτηριστικά που αποδίδουμε συνήθως σε ανθρώπινη συμπεριφορά, όπως η επίλυση προβλημάτων, η αντίληψη μέσω της όρασης, η μάθηση, η εξαγωγή συμπερασμάτων, η κατανόηση φυσικής γλώσσας, κλπ.

Μερικοί άνθρωποι θεωρούν ότι η ΑΙ αποτελεί κίνδυνο για την ανθρωπότητα αν προχωρήσει ανεξέλεγκτα. Ένας από αυτούς ήταν και ο Βρετανός, φυσικός, συγγραφέας και κοσμολόγος Stephen William Hawking (1942–2018), ο οποίος όπως είχε ισχυριστεί "η τεχνητή νοημοσύνη θα μπορούσε να είναι το τελευταίο επίτευγμα της ανθρωπότητας".¹³ Οι προσπάθειες δημιουργίας τεχνητής νοημοσύνης έχουν βιώσει πολλές αποτυχίες, όπως η έκθεση ALPAC του 1966, η εγκατάλειψη των perceptrons το 1970, η έκθεση Lighthill του 1973, ο δεύτερος χειμώνας του 1987-1993 και η κατάρρευση της αγοράς μηχανών Lisp το 1987. Στον εικοστό πρώτο αιώνα, οι τεχνικές ΑΙ, τόσο αυτές που χρησιμοποιούν συμβολική προσέγγιση όσο και οι "υποσυμβολικές", έχουν βιώσει μια αναζωπύρωση μετά από ταυτόχρονες εξελίξεις στην δύναμη του υπολογιστή, τα μεγέθη των συστημάτων εκπαίδευσης και τη θεωρητική κατανόηση. Οι πρόσφατες εξελίξεις στην τεχνολογία της πνευματικής ιδιοκτησίας και ειδικότερα στην εκμάθηση μηχανών, συνέβαλαν στην ανάπτυξη των αυτόνομων πραγμάτων όπως τα αεροσκάφη, τα αυτοκίνητα κτλ.

2.4. Υπολογιστική Νοημοσύνη (Computational Intelligence - CI)

Η Υπολογιστική Νοημοσύνη αρχικά ορίστηκε ως το σύνολο τριών τεχνολογιών που περιλαμβάνουν τα νευρωνικά δίκτυα, τα ασαφή συστήματα και τον εξελικτικό υπολογισμό.

⁹ Haugeland, 1985.

¹⁰ Rich and Knight, 1991.

¹¹ Winston, 1992.

¹² Luger and Stubblefield, 1993.

¹³ BetaNews, 21 Oktwmbriou 2016.

Αργότερα, προτάθηκαν επιπλέον τεχνολογίες όπως συστήματα στήριξης αποφάσεων, μηχανικής μάθησης, εξόρυξης δεδομένων, και ποικίλες συνέργιες αυτών. Πρόσφατα δημοσιεύματα τοποθετούν την Υπολογιστική Νοημοσύνη στον πυρήνα τεχνολογιών αιχμής που έχουν να κάνουν με την επεξεργασία τεράστιου όγκου δεδομένων, την αλληλεπίδραση ανθρώπου-υπολογιστή, το Διαδίκτυο πραγμάτων, κλπ.¹⁴ Επίσης, ένας άλλος ορισμός χαρακτηρίζει την Υπολογιστική Νοημοσύνη σαν ένα σύνολο αλγόριθμων που βασίζονται σε φυσικές διαδικασίες, οι οποίες αναπτύχθηκαν κατά τη διάρκεια των τριών τελευταίων δεκαετιών και απέκτησε πρακτικό ενδιαφέρον με την εμφάνιση των παράλληλων ηλεκτρονικών υπολογιστών.

Η υπολογιστική νοημοσύνη ακούγεται παντού και συχνά, αλλά σημαίνει διαφορετικά πράγματα σε διαφορετικούς ανθρώπους. Το όνομα "Computational Intelligence" έχει χρησιμοποιηθεί για πάνω από 20 χρόνια, αν και μόνο πρόσφατα έχει αποκτήσει ευρεία δημοτικότητα και κάπως διαφορετική ερμηνεία. Πρόκειται για ένα πολύ συγκεχυμένο ζήτημα, το οποίο συζητήθηκε έντονα, αλλά δεν υπήρξε συναίνεση. Χαρακτηρίζεται κυρίως από ορισμένες σημαντικές ιδιότητες, που είναι χαρακτηριστικά συστημάτων που χρησιμοποιούν τεχνικές μηχανικής μάθησης, όπως: προσαρμοστικότητα (adaptation), αυτό-οργάνωση (self-organization), μάθηση-εξέλιξη (learning-evolution).

Μέχρι τώρα, δεν υπάρχει κανένας κοινώς αποδεκτός ορισμός της υπολογιστικής νοημοσύνης. Γενικά, η υπολογιστική νοημοσύνη είναι ένα σύνολο εμπνευσμένο από τη φύση με υπολογιστικές μεθοδολογίες και προσεγγίσεις για την αντιμετώπιση πολύπλοκων προβλημάτων του πραγματικού κόσμου, στα οποία τα μαθηματικά ή τα παραδοσιακά μοντέλα μπορεί να είναι άχρηστα για διάφορους λόγους. Αυτά, για παράδειγμα, μπορεί να είναι προβλήματα που οι διαδικασίες να είναι υπερβολικά πολύπλοκες για την μαθηματική λογική ή μπορεί να υπάρχουν ορισμένες αβεβαιότητες. Πράγματι, πολλά προβλήματα δεν μπορούν να μεταφραστούν σε δυαδική γλώσσα (μοναδικές τιμές 0 και 1) για τους υπολογιστές, ώστε να τα επεξεργαστούν. Η υπολογιστική νοημοσύνη επομένως παρέχει τις λύσεις σε διαδικασίες υπερβολικά πολύπλοκες για τη μαθηματική λογική ή σε προβλήματα που μπορεί να υπάρχουν ορισμένες αβεβαιότητες.



Εικόνα 1:
Αλληλεπίδραση ανθρώπου - υπολογιστή σε 0-1

¹⁴ M. M. Polycarpou, «Computational intelligence in the undergraduate curriculum», *IEEE Computational Intelligence Magazine*, vol. 8, no. 2, p. 3, Μάιος 2013

Οι μέθοδοι που χρησιμοποιεί είναι κοντά στην ανθρώπινη λογική, δηλαδή χρησιμοποιεί μη ακριβή και μη πλήρη γνώση, και είναι σε θέση να παραγάγει τις ενέργειες ελέγχου με έναν προσαρμοστικό τρόπο. Το CI επομένως χρησιμοποιεί έναν συνδυασμό των πέντε κύριων συμπληρωματικών τεχνικών: την ασαφή λογική (fuzzy logic) που επιτρέπει στον υπολογιστή να καταλάβει τη φυσική γλώσσα, τα τεχνητά νευρικά δίκτυα (artificial neural networks) που επιτρέπουν στο σύστημα να μάθει τα εμπειρικά στοιχεία λειτουργώντας τα, τον εξελικτικό προγραμματισμό (evolutionary computing) ο οποίος είναι βασισμένος στη διαδικασία της φυσικής επιλογής, τη θεωρία μάθησης και τις πιθανολογικές μεθόδους που βοηθούν στην ανακρίβεια.

Η Υπολογιστική Νοημοσύνη είναι ένας τρόπος μίμησης των ανθρώπινων όντων. Πράγματι, το χαρακτηριστικό της "νοημοσύνης" αποδίδεται συνήθως στον άνθρωπο. Πρόσφατα, πολλά προϊόντα και αντικείμενα ισχυρίζονται, επίσης, ότι είναι "έξυπνα", ένα χαρακτηριστικό το οποίο συνδέεται άμεσα με την λήψη αποφάσεων και συλλογισμού.

Ο όρος Υπολογιστική Νοημοσύνη αφορά κάθε είδος μη βιολογικής νοημοσύνης που μπορεί να εκδηλωθεί με υπολογιστικές διαδικασίες. Η υπολογιστική νοημοσύνη έχει σχεδιαστεί για να αντιγράψει τις πτυχές της βιολογικής νοημοσύνης. Πολλές φορές προβλήματα καθημερινότητας δεν μπορούμε να τα λύσουμε λόγω αβεβαιότητας, άγχους, παρουσίας θορύβου κ.λπ. Η υπολογιστική νοημοσύνη μπορεί να μας βοηθήσει σε αυτό, αν και δεν είναι μαθηματικά προβλήματα, χρησιμοποιώντας συγκεκριμένες μεθοδολογίες. Οι μέθοδοι της Υπολογιστικής Νοημοσύνης, που βασίζονται σε παραδείγματα από τη φύση, έχουν οδηγήσει στην ανάπτυξη μιας μεγάλης ποικιλίας αλγορίθμων και μεθόδων, όπως τα Τεχνητά Νευρωνικά Δίκτυα, τους Γενετικούς Αλγόριθμους (ΓΑ) και γενικότερα τους Εξελικτικούς Αλγόριθμους (ΕΑ), τα Ασαφή Συστήματα (ΑΣ), την Ευφυΐα Σμήνους (ΕΣ), τα συστήματα ταξινόμησης, τις εξελικτικές στρατηγικές, τις μεθόδους τοπικής αναζήτησης κτλ., που οδηγούν σε αποδοτικές και ικανοποιητικές λύσεις. Μάλιστα τελευταία, έχουν αναπτυχθεί καινούριες μέθοδοι οι οποίες αξιοποιούν τα πλεονεκτήματα δύο ή περισσότερων από τις παραπάνω μεθόδους, εξαλείφοντας έτσι τα μειονεκτήματα που έχει η κάθε μια ξεχωριστά.¹⁵

Η υπολογιστική νοημοσύνη μπορεί να διακριθεί στις παρακάτω κατηγορίες: Τα νευρωνικά δίκτυα, τον εξελικτικό υπολογισμό, την ασαφή λογική και τους γενετικούς αλγόριθμους που συμπεριλαμβάνουν συστήματα επίλυσης προβλημάτων βασισμένων στις αρχές της Γενετικής Εξέλιξης και της Κληρονομικότητας. Τα συστήματα αυτά αντιγράφουν τον τρόπο με τον οποίο αναπαράγονται και μεταλλάσσονται τα χρωμοσώματα των κυττάρων των ζωντανών οργανισμών και δημιουργούν συνεχώς νέους πληθυσμούς πιθανών λύσεων, χρησιμοποιώντας κομμάτια από την προηγούμενη γενιά και την καινούρια.

2.4.1. Ανάλυση των κύριων συμπληρωματικών τεχνικών της υπολογιστικής νοημοσύνης.

- **Ασαφής λογική (fuzzy logic)**

Όπως εξηγήθηκε προηγουμένως, η ασαφής λογική, μία από τις βασικές αρχές του CI, συνίσταται σε μετρήσεις και μοντελοποίηση διαδικασιών που γίνονται για πολύπλοκες διαδικασίες της πραγματικής ζωής. Μπορεί να αντιμετωπίσει την ατέλεια και, κυρίως, την άγνοια των δεδομένων σε ένα μοντέλο διαδικασίας, αντίθετα με την Τεχνητή Νοημοσύνη, η οποία απαιτεί ακριβή γνώση.

¹⁵ © Καβαλιεράτου, Copyright ICSD :: 2008 - 2018 Πανεπιστήμιο Αιγαίου.
http://www.icsd.aegean.gr/lecturers/kavallieratou/NN&EP_files/ci_1.pdf.

Αυτή η τεχνική τείνει να εφαρμόζεται σε ένα ευρύ φάσμα τομέων όπως ο έλεγχος, η επεξεργασία εικόνων και η λήψη αποφάσεων. Είναι όμως επίσης γνωστό στον τομέα των οικιακών συσκευών με πλυντήρια, φούρνους μικροκυμάτων κλπ. Μπορούμε να το συναντήσουμε και όταν χρησιμοποιούμε μια βιντεοκάμερα, όπου βοηθάει στη σταθεροποίηση της εικόνας. Αλλά και σε άλλους τομείς όπως η ιατρική διάγνωση και η διαπραγμάτευση συναλλαγματικών ισοτιμιών.

Η ασαφής λογική είναι κυρίως χρήσιμη για την προσεγγιστική συλλογιστική και δεν έχει ικανότητες μάθησης. Ένα απαιτούμενο προσόν που χρειάζονται τα ανθρώπινα όντα που τους επιτρέπει να βελτιώσουν τον εαυτό τους μαθαίνοντας από τα προηγούμενα λάθη τους.

- **Νευρωνικά δίκτυα**

Αυτός είναι ο λόγος για τον οποίο οι εμπειρογνώμονες του CI εργάζονται στην ανάπτυξη τεχνητών νευρωνικών δικτύων με βάση τα βιολογικά δίκτυα, τα οποία μπορούν να οριστούν από 3 κύρια στοιχεία: το κύτταρο-σώμα, που επεξεργάζεται τις πληροφορίες, τον άξονα, που είναι μια συσκευή που επιτρέπει τη διεξαγωγή του σήματος και την συνάφεια, η οποία ελέγχει τα σήματα. Επομένως, τα τεχνητά νευρωνικά δίκτυα εξαντλούνται από καταναμημένα συστήματα επεξεργασίας πληροφοριών, επιτρέποντας τη διαδικασία και τη μάθηση από βιωματικά δεδομένα. Εργαζόμενα σαν ανθρώπινα όντα, η ανοχή σφάλματος είναι επίσης ένα από τα βασικά πλεονεκτήματα αυτής της αρχής.

Όσον αφορά τις εφαρμογές τους, τα νευρωνικά δίκτυα μπορούν να ταξινομηθούν σε πέντε ομάδες: ανάλυση και ταξινόμηση δεδομένων, συσσωρευτική μνήμη, δημιουργία ομάδων, δημιουργία μοτίβων και έλεγχος. Γενικά, αυτή η μέθοδος στοχεύει στην ανάλυση και ταξινόμηση των ιατρικών δεδομένων, στη διερεύνηση της αυθεντικότητας της ταυτότητας του προσώπου και, κυρίως, στην αντιμετώπιση των μη γραμμικών στοιχείων ενός συστήματος προκειμένου να το ελέγξει. Επιπλέον, οι τεχνικές των νευρωνικών δικτύων μοιράζονται με τις ασαφείς λογικές το πλεονέκτημα της ενεργοποίησης ομαδοποίησης δεδομένων.

- **Εξελικτικός προγραμματισμός (evolutionary computing)**

Με βάση τη διαδικασία φυσικής επιλογής που εισήχθη αρχικά από τον Charles Robert Darwin,¹⁶ ο εξελικτικός υπολογισμός συνίσταται στην κεφαλαιοποίηση της δύναμης της φυσικής εξέλιξης για την εμφάνιση νέων τεχνητών εξελικτικών μεθοδολογιών.¹⁷ Περιλαμβάνει επίσης άλλους τομείς, όπως την εξελικτική στρατηγική και τους εξελικτικούς αλγορίθμους, που αντιμετωπίζονται ως μέθοδοι επίλυσης προβλημάτων.

Οι βασικές εφαρμογές αυτής της αρχής καλύπτουν τομείς όπως η βελτιστοποίηση ενός στόχου και η βελτιστοποίηση πολλαπλών στόχων, στους οποίους οι παραδοσιακές μαθηματικές τεχνικές δεν αρκούν πια για να εφαρμοστούν σε ένα ευρύ φάσμα προβλημάτων όπως η Ανάλυση του DNA, διάφορα προβλήματα προγραμματισμού, κτλ.

- **Θεωρία Μάθησης**

Ακόμα ψάχνουμε για έναν τρόπο «σκέψης» που να πλησιάζει τον ανθρώπινο. Η θεωρία μάθησης είναι μία από τις κύριες προσεγγίσεις της υπολογιστικής νοημοσύνης. Στην ψυχολογία, η μάθηση είναι η διαδικασία συγκέντρωσης γνωστικών, συναισθηματικών και περιβαλλοντικών επιδράσεων και εμπειριών για την απόκτηση, ενίσχυση ή αλλαγή γνώσεων, δεξιοτήτων, αξιών και παγκόσμιων απόψεων.¹⁸ Οι θεωρίες μάθησης στη συνέχεια βοηθούν

¹⁶ Κάρολος Δαρβίνος, 1809-1882, Άγγλος φυσικός, γεωλόγος και βιολόγος.

¹⁷ De Jong, «Evolutionary Computation: A Unified Approach», MIT Press. ISBN 9780262041942, K. 2006.

¹⁸ Ormrod, 1995, Illeris, 2004.

στην κατανόηση του τρόπου επεξεργασίας αυτών των αποτελεσμάτων και εμπειριών και, στη συνέχεια, βοηθούν στην πραγματοποίηση προβλέψεων βάσει προηγούμενης εμπειρίας.

- **Πιθανολογικές Μέθοδοι - Πιθανότητες**

Ως ένα από τα βασικά στοιχεία της ασαφούς λογικής, οι πιθανοτικές μέθοδοι που εισήχθησαν αρχικά από τον Ούγγρο μαθηματικό Paul Erdos και τον Joel Spencer,¹⁹ αποσκοπούν στην εκτίμηση των αποτελεσμάτων ενός συστήματος ευφών υπολογισμών, το οποίο προσδιορίζεται κυρίως από τυχαία σειρά. Επομένως, οι πιθανοτικές μέθοδοι αναδεικνύουν τις πιθανές λύσεις σε ένα πρόβλημα συλλογισμού, βασισμένο σε προηγούμενες γνώσεις.

2.4.2. Απόψεις επιστημόνων για την υπολογιστική νοημοσύνη.

Ο όρος χρησιμοποιήθηκε πρώτη φορά από τον καθηγητή J. Bezdek (1992), στο επιστημονικό περιοδικό International Journal of Approximate Reasoning. Σε άρθρο του που αφορούσε τα τεχνητά νευρωνικά δίκτυα, ο Bezdek εισήγαγε τον όρο ABCs για να αποσαφηνίσει τα παρακάτω: A= Artificial: Non-Biological (Man-Made), B= Biological: Physical + Chemical = Organic και C= Computational: Mathematics + Man-Made Machines.

	Input	Complexity →			Level		
↑ Complexity	Sensory Data + Human Knowledge	BNN	C	BPR	C	BI	B ~ Organic
	Sensor Data + "Knowledge Tidbits"	U		U		U	
	ANN	C	APR	C	AI	A ~ Symbolic	
	Sensor Data + Computation	U		U		U	
		CNN	C	CPR	C	CI	C ~ Numeric

Εικόνα 2:

The ABCs: Neural Networks, Pattern Recognition and Intelligence (Bezdek, 1992)

(όπου A: artificial, B: biological, C: computational, NN: neural / neural like, PR: pattern recognition)

Επίσης, το 1993 διαφοροποίησε την CI από την AI (Artificial Intelligence – Τεχνητή Νοημοσύνη), με το επιχείρημα ότι το πρώτο βασίζεται σε εύκαμπτους υπολογιστικούς μεθόδους, ενώ η AI βασίζεται στο σκληρό υπολογισμό αυτών. Τέλος, το 1994, έδωσε τον πρώτο σαφή ορισμό για την ΥΝ: "ένα σύστημα καλείται υπολογιστικά ευφές εάν ασχολείται με τα δεδομένα χαμηλού επιπέδου, όπως αριθμητικά δεδομένα, εάν έχει ένα στοιχείο αναγνώρισης προτύπων και εάν δεν χρησιμοποιεί τη γνώση με τόση ακρίβεια και πληρότητα όσο η τεχνητή νοημοσύνη".²⁰

¹⁹ Siddique, Nazmul; Adeli, Hojjat (2013), «Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing». John Wiley & Sons. ISBN 978-1-118-53481-6.

²⁰ James. C. Bezdek, «What is computational intelligence? In: Computational Intelligence Imitating».

Ένας αποδεκτός ορισμός για την ΥΝ είναι και ο ακόλουθος: "Υπολογιστική Νοημοσύνη είναι ο επιστημονικός χώρος που προσφέρει τις τεχνικές για την επίλυση δύσκολων προβλημάτων, με τη μηχανή να μιμείται απλώς, βιολογικές διεργασίες, χωρίς να είναι απαραίτητο να επιδεικνύει γενική νοημοσύνη".²¹ Από την άλλη, επιστήμονες ισχυρίζονται ότι η Υπολογιστική Νοημοσύνη είναι το σύνολο των αλγορίθμων που βασίζονται σε φυσικές διαδικασίες και αναπτύχθηκαν κατά τη διάρκεια των τριών τελευταίων δεκαετιών και απόκτησε πρακτικό ενδιαφέρον με την εμφάνιση των παράλληλων ηλεκτρονικών υπολογιστών. Μία Τρίτη άποψη είναι ότι η CI είναι ένας τρόπος μίμησης των ανθρώπινων όντων. Πράγματι, το χαρακτηριστικό της "νοημοσύνης" αποδίδεται συνήθως στον άνθρωπο. Πρόσφατα, πολλά προϊόντα και αντικείμενα ισχυρίζονται, επίσης, ότι είναι "έξυπνα", ένα χαρακτηριστικό το οποίο συνδέεται άμεσα με την λήψη αποφάσεων και συλλογισμού.

Σύμφωνα με ένα άρθρο του πανεπιστημίου του Άμστερνταμ, οι τομείς που καλύπτονται από τον όρο υπολογιστική ευφυΐα είναι επίσης γνωστές με το όνομα soft computing. Και πάλι, σύμφωνα με την επιστημονική λαογραφία, αυτό το όνομα επελέγη για να υποδείξει τη διαφορά μεταξύ του soft computing και της έρευνας των λειτουργιών, γνωστή και ως hard computing. Οι δύο αυτοί τομείς συνδέονται με τους τομείς προβλημάτων στους οποίους εφαρμόζονται, αλλά ενώ οι αλγόριθμοι έρευνας διεξάγονται συνήθως με σκληρές και αυστηρές συνθήκες σχετικά με το εύρος εφαρμογής και τις αποδεδειγμένες εγγυήσεις για μια βέλτιστη λύση, το soft computing δεν θέτει προϋποθέσεις για το πρόβλημα όπως επίσης και δεν παρέχει καμία εγγύηση για την επιτυχία, μια ανεπάρκεια που αντισταθμίζεται από την ευρωστία των μεθόδων.

Οι συγγραφείς David Poole (καθηγητής στον τομέα της επιστήμης των υπολογιστών, στο πανεπιστήμιο της Κολούμπια), και ο συνάδελφός του Alan Mackworth, στο βιβλίο που δημοσίευσαν ισχυρίζονται ότι η υπολογιστική νοημοσύνη είναι η μελέτη του σχεδιασμού ευφυών μέσων όπως χαρακτηριστικά αναφέρουν: "intelligent agent (IA)".²²

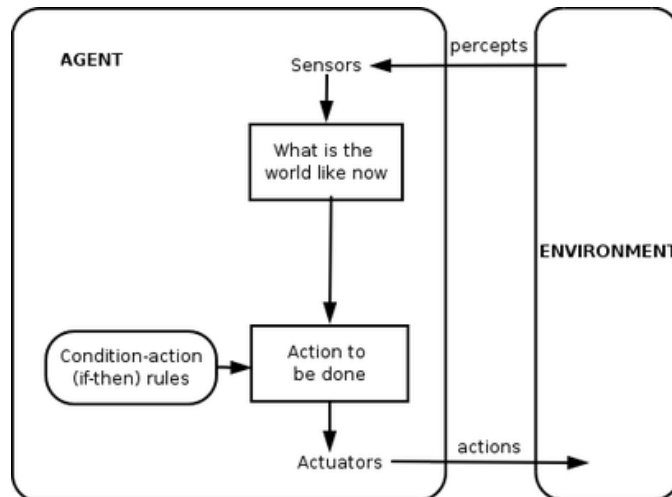
Ένας intelligent agent είναι αυτός που αντιλαμβάνεται και δρα στο περιβάλλον, ένα μέσων, του οποίου οι αποφάσεις σχετικά με τις πράξεις του μπορούν να εξηγηθούν από άποψη υπολογισμών. Αυτός ο υπολογισμός μπορεί να λάβει πολλές μορφές. Στους υπολογιστές πραγματοποιείται στο "υλικό-hardware". Πιο συγκεκριμένα σύμφωνα με τον Franklin, Graesser, Stan και Art (1996): "Αν ένας παράγοντας πράττει έτσι ώστε να μεγιστοποιήσει την προσδοκώμενη αξία του μέτρου απόδοσης, που βασίζεται σε παλαιότερη εμπειρία και γνώση, τότε είναι νοήμων."

Αν και υπάρχουν ορισμένοι "agents" (παράγοντες) που είναι αναμφισβήτητα μη υπολογιστικοί, όπως ο άνεμος και η βροχή που διαβρώνουν ένα τοπίο, είναι ανοιχτό το ερώτημα εάν όλοι οι "intelligent agents" είναι υπολογιστικοί. Το μόνο σίγουρο είναι ότι έχουν την δυνατότητα να ενεργούν έξυπνα. Είναι ευέλικτοι στα μεταβαλλόμενα περιβάλλοντα και τους μεταβαλλόμενους στόχους. Μαθαίνουν από την εμπειρία και κάνουν της καλύτερες επιλογές λαμβάνοντας υπ' όψιν περιορισμούς και υπολογισμούς.

Ο κεντρικός επιστημονικός στόχος της υπολογιστικής νοημοσύνης είναι να κατανοηθούν οι αρχές που καθιστούν δυνατή την έξυπνη συμπεριφορά, σε τεχνητά ή φυσικά συστήματα. Η βασική υπόθεση είναι ότι ο συλλογισμός και ο βασικός στόχος του μηχανικού είναι να προσδιορίσει μεθόδους για το σχεδιασμό χρήσιμων και έξυπνων αντικειμένων.

²¹ Σωκράτης Κ. Κάτσικας Αθήνα, Σεπτέμβρης 2015.

²² Foundations of Computational Agents, Cambridge University Press, 2010.



Εικόνα 3:
Simple reflex agent, based on Artificial Intelligence: A Modern Approach
 (20 January 2006, Utkarshraj Atmaram)

Παρότι δεν υπάρχει κοινά αποδεκτός ορισμός, η Υπολογιστική Νοημοσύνη μπορεί να θεωρηθεί ως η μελέτη των μηχανισμών που κάνουν δυνατή την ευφυή συμπεριφορά ενός συστήματος σε ένα πολύπλοκο και μεταβαλλόμενο περιβάλλον. Υπολογιστική Νοημοσύνη είναι ένα σύνολο εμπνευσμένων από την φύση υπολογιστικών μεθοδολογιών και προσεγγίσεων για την αντιμετώπιση πολύπλοκων προβλημάτων του πραγματικού κόσμου, στον οποίο παραδοσιακές μεθοδολογίες και προσεγγίσεις είναι αναποτελεσματικές ή ανέφικτες.

Ο όρος Υπολογιστική Νοημοσύνη, υποδηλώνει κάθε είδος νοημοσύνης που μπορεί να εκδηλωθεί με υπολογιστικές διαδικασίες. Στην υπολογιστική νοημοσύνη, η βασική ιδέα-προσέγγιση είναι αντί να αναπαριστούμε την γνώση σε ένα υψηλό επίπεδο, αντικαθιστούμε τα μοντέλα αυτά με αλγόριθμους που στηρίζονται στην αριθμητική επεξεργασία της πληροφορίας (δηλαδή υπολογιστικά). Έτσι, αντί να έχουμε ένα αναλυτικό μαθηματικό μοντέλο για την εφαρμογή, κατασκευάζουμε έναν αλγόριθμο ο οποίος αναλύει και μαθαίνει από δεδομένα που ήδη υπάρχουν και περιγράφουν την εφαρμογή.²³

Με σκοπό να βρεθεί συναίνεση και για την ενημέρωση πάνω στην εξέλιξη της υπολογιστικής νοημοσύνης, ιδρύθηκε το Παγκόσμιο Συνέδριο για την Υπολογιστική Νοημοσύνη. Μέχρι το 2006 διεξάγονταν κάθε τέσσερα χρόνια (1994 Ορλάντο, 1998 Anchorage Alaska, 2002 Χονολουλού, 2006 Vancouver Canada), και από τότε κάθε 2 χρόνια (2008 Hong Kong, 2010 Μπαρτσελόνα, 2012 Brisbane Αυστραλία, 2014 Πεκίνο, 2016 Vancouver και αναμένεται το επόμενο για το 2018 στο Rio De Janeiro). Αποτελείται από τρία βασικά τμήματα: το Διεθνές Συνέδριο IEEE για την Εξέλιξη Υπολογιστών, την Ασαφή Υπολογιστική και τα νευρωνικά δίκτυα.

²³ Χρίστος Ν. Σχίζας, «Εισαγωγή στην Υπολογιστική Νοημοσύνη», Συστήματα Υπολογιστικής Νοημοσύνης – Σημειώσεις Διαλέξεων, Τμήμα Πληροφορικής, Πανεπιστήμιο Κύπρου 2010.

2.5. Αποσαφήνιση των όρων Υπολογιστική Νοημοσύνη και Τεχνητή Νοημοσύνη.

Με την πάροδο του χρόνου, διάφοροι κλάδοι της υπολογιστικής Νοημοσύνης εξελίσσονται σε πολλές κατευθύνσεις, που συχνά είναι αρκετά μακριά από τις αρχικές ρίζες και τις εμπνεύσεις. Οι Νέες κοινότητες που σχηματίζονται, πρέπει να καθιερώσουν την ταυτότητά τους καθορίζοντας τα σύνορα που τις διακρίνουν από άλλες επιστημονικές κοινότητες.

Η Τεχνητή Νοημοσύνη (AI) ήταν η πρώτη μεγάλη επιστημονική κοινότητα που ιδρύθηκε ήδη στα μέσα της δεκαετίας του 1950 και αφορούσε προβλήματα που απαιτούν νοημοσύνη για να λυθούν και εκδηλώνεται από μηχανές ή λογισμικό. Οι κύριοι ερευνητές και βιβλία AI ορίζουν το πεδίο ως "μελέτη και σχεδιασμό ευφύων μέσων", όπου ένα ευφές μέσο (intelligent agent), όπως αναφέραμε παραπάνω, είναι ένα σύστημα που αντιλαμβάνεται το περιβάλλον του και αναλαμβάνει δράσεις που μεγιστοποιούν τις πιθανότητες επιτυχίας σε κάποιο στόχο του.

Η υπολογιστική νοημοσύνη είναι ένα σύνολο από υπολογιστικές μεθοδολογίες και προσεγγίσεις για την αντιμετώπιση πολύπλοκων πραγματικών προβλημάτων στα οποία οι παραδοσιακές προσεγγίσεις, η μοντελοποίηση των πρώτων αρχών ή η ρητή στατιστική μοντελοποίηση είναι αναποτελεσματικές ή μη εφικτές. Πολλά από αυτά τα προβλήματα της πραγματικής ζωής δεν θεωρούνται μαθηματικά προβλήματα, αλλά η φύση παρέχει πολλά αντίθετα δείγματα βιολογικών συστημάτων που παρουσιάζουν την απαιτούμενη λειτουργία, πρακτικά.

Η Τεχνητή Νοημοσύνη και η Υπολογιστική Νοημοσύνη επιδιώκουν έναν παρόμοιο μακροπρόθεσμο στόχο: να φτάσουν στη γενική νοημοσύνη, η οποία είναι η νοημοσύνη μιας μηχανής που θα μπορούσε να εκτελέσει οποιαδήποτε πνευματική εργασία που μπορεί να κάνει ένας άνθρωπος. Υπάρχει όμως σαφής διαφορά μεταξύ τους.

Σύμφωνα με τον Bezdek (1994), η Υπολογιστική Νοημοσύνη είναι ένα υποσύνολο της Τεχνητής Νοημοσύνης. Ο R. Marks (1993) ανέφερε για τη σχέση YN και TN ότι "Παρόλο που επιδιώκουν παρόμοιους στόχους, η Υπολογιστική Νοημοσύνη αναδύθηκε ως ένας ανεξάρτητος κλάδος, του οποίου το ερευνητικό πεδίο είναι κάπως διαφορετικό από αυτό της Τεχνητής Νοημοσύνης". Επίσης συμφώνησε με τον Bezdek ότι η υπολογιστική νοημοσύνη βασίζεται σε soft computing μεθόδους, ενώ η τεχνητή νοημοσύνη βασίζεται σε hard computing μεθόδους.

Η συμβατική έρευνα του AI επικεντρώνεται στις προσπάθειες να μιμηθεί την ανθρώπινη νοημοσύνη μέσω χειραγώγησης συμβόλων και συμβολικά δομημένων βάσεων γνώσεων. Αυτή η προσέγγιση περιορίζει τις καταστάσεις στις οποίες μπορεί να εφαρμοστεί συμβατικό AI. Ο Lotfi Zadeh δήλωσε ότι "διαθέτουμε επίσης υπολογιστικά εργαλεία τα οποία είναι πολύ πιο αποτελεσματικά στη σύλληψη και το σχεδιασμό έξυπνων συστημάτων από τις μεθόδους βασισμένες στην λογική που βασίζονται στην λογική που αποτελούν τον πυρήνα του παραδοσιακού AI". Αυτές οι τεχνικές, οι οποίες περιλαμβάνουν την ασαφή λογική, έχουν γίνει γνωστές ως soft computing. Αυτές οι συχνά βιολογικά εμπνευσμένες μέθοδοι βρίσκονται σε αντίθεση με το συμβατικό AI και αντισταθμίζουν τις αδυναμίες του συμβολισμού. Αυτές οι δύο μεθοδολογίες έχουν επίσης επισημανθεί ως neats vs. scruffies, ενώ οι neats υπογραμμίζουν τη χρήση της λογικής και της επίσημης αναπαράστασης της γνώσης, ενώ οι scruffies υιοθετούν μια προσέγγιση ηθικής

Οι hard computing τεχνικές υπολογιστών, λειτουργούν ακολουθώντας τη δυαδική λογική που βασίζεται μόνο σε δύο τιμές (τα Booleans true ή false, 0 ή 1), στα οποία βασίζονται οι σύγχρονοι υπολογιστές. Ένα πρόβλημα με αυτή τη λογική είναι ότι η φυσική μας γλώσσα δεν μπορεί πάντα να μεταφραστεί εύκολα σε απόλυτους όρους των 0 και 1. Οι

soft τεχνικές υπολογιστικής είναι βασισμένες σε ασαφείς λογικές²⁴. Πολύ πιο κοντά στον τρόπο με τον οποίο ο ανθρώπινος εγκέφαλος λειτουργεί, συγκεντρώνοντας δεδομένα σε μερικές αλήθειες (Crisp / ασαφή συστήματα), αυτή η λογική είναι μία από τις κύριες αποκλειστικές πτυχές της υπολογιστικής νοημοσύνης.

Αν και οι δύο όροι, και ο όρος Υπολογιστική Νοημοσύνη αλλά και ο όρος Τεχνητή Νοημοσύνη, υποδηλώνουν κάθε είδος νοημοσύνης που μπορεί να εκδηλωθεί με υπολογιστικές διαδικασίες, ο όρος της τεχνητής νοημοσύνης θεωρείται από κάποιους επιστήμονες ότι είναι πιο γενικός, υπό την έννοια ότι υποδηλώνει κάθε είδος νοημοσύνης που δεν είναι βιολογική. Υπάρχουν δύο προσεγγίσεις Τεχνητής νοημοσύνης. Η ασθενής προσέγγιση, στην οποία πιστεύεται ότι η ανθρώπινη γνώση/νόηση ή ειδικές εκφάνσεις αυτής, μπορεί να προσομοιαστούν σε υπολογιστικά συστήματα, και η ισχυρή προσέγγιση, στην οποία πιστεύεται ότι η ανθρώπινη γνώση/νόηση μπορεί να εκδηλωθεί αυτόματα σε υπολογιστικά συστήματα. Η Τεχνητή νοημοσύνη χρησιμοποιεί πολύπλοκα ηλεκτρονικά συστήματα για να υλοποιήσει τα μοντέλα της. Στην υπολογιστική νοημοσύνη, η βασική ιδέα-προσέγγιση είναι αντί να αναπαριστούμε την γνώση σε ένα υψηλό επίπεδο, αντικαθιστούμε τα μοντέλα αυτά με αλγόριθμους που στηρίζονται στην αριθμητική επεξεργασία της πληροφορίας (δηλαδή υπολογιστικά). Έτσι, αντί να έχουμε ένα αναλυτικό μαθηματικό μοντέλο για την εφαρμογή, κατασκευάζουμε έναν αλγόριθμο ο οποίος αναλύει και μαθαίνει από δεδομένα που ήδη υπάρχουν και περιγράφουν την εφαρμογή.²⁵ Η υπολογιστική νοημοσύνη θεωρείται ένα «μήνυμα», που σύμφωνα με την επιστημονική λαογραφία επιλέχθηκε για να υποδηλώνει τη σύνδεση και τη διαφορά με την τεχνητή νοημοσύνη.

Πέρα από όλες αυτές τις διαφορές που αναφέραμε παραπάνω, ορισμένες τεχνικές υπολογιστικής νοημοσύνης συχνά θεωρούνται ως τεχνικές τεχνητής νοημοσύνης (π.χ. γενετικοί αλγόριθμοι ή νευρωνικά δίκτυα), αλλά ακόμα και εδώ υπάρχει μια σαφής διαφορά μεταξύ αυτών των τεχνικών και των παραδοσιακών τεχνικών τεχνητής νοημοσύνης. Και αυτή η διαφορά είναι με βάση τη λογική. Γενικά, οι τυπικές τεχνικές τεχνητής νοημοσύνης είναι από πάνω προς τα κάτω όπου, π.χ., η δομή των μοντέλων, των λύσεων κ.λπ. Οι τεχνικές υπολογιστικής νοημοσύνης είναι γενικά από τη βάση προς την κορυφή, όπου η τάξη και η δομή εξέρχονται από μια αδόμενη αρχή.

Πολλοί είναι αυτοί βέβαιοι, που θεωρούν ότι τα τελευταία χρόνια οι άνθρωποι άρχισαν να χρησιμοποιούν τους όρους «μηχανική νοημοσύνη» και «υπολογιστική νοημοσύνη» αντί για «τεχνητή νοημοσύνη» επειδή για τους επαγγελματίες, ο τελευταίος όρος, έχει συσχετιστεί πάρα πολύ με τις παλιές ιδέες των μηχανών και της συμβολικής λογικής. Εκτός από αυτούς που ισχυρίζονται ότι η υπολογιστική νοημοσύνη είναι ένα κομμάτι της τεχνητής, η προσέγγιση που ακολουθήθηκε από τα περιοδικά και από τους συγγραφείς βιβλίων είναι να θεωρούν την υπολογιστική νοημοσύνη ως μία ομπρέλα κάτω από την οποία όλο και περισσότερες μέθοδοι σιγά-σιγά προστίθενται. Συνεπώς, ένας καλός ορισμός αυτών των πεδίων είναι αδύνατος, επειδή διαφορετικοί άνθρωποι προσθέτουν ή αποκλείουν διαφορετικές μεθόδους κάτω από την ίδια κλάση Υπολογιστικής Νοημοσύνης και τεχνητής νοημοσύνης.

Η σύγχυση σχετικά με το όνομα της υπολογιστικής νοημοσύνης, μπορεί εν μέρει να αποδοθεί στην σύγχυση που υπάρχει όσο αφορά τον σκοπό του πεδίου αυτού. Ο στόχος της

24 «Artificial Intelligence, Computational Intelligence, SoftComputing, Natural Computation - what's the difference? - ANDATA». www.andata.at. Retrieved, Νοέμβριος 2015.

25 Χρίστος Ν. Σχίζας, «Εισαγωγή στην Υπολογιστική Νοημοσύνη», Συστήματα Υπολογιστικής Νοημοσύνης – Σημειώσεις Διαλέξεων, Τμήμα Πληροφορικής, Πανεπιστήμιο Κύπρου 2010.

υπολογιστικής νοημοσύνης είναι να σχεδιαστούν, να κατασκευαστούν και να γίνουν πειράματα με υπολογιστικά συστήματα που εκτελούν εργασίες που συνήθως θεωρούνται έξυπνες. Η υπολογιστική νοημοσύνη περιλαμβάνει επαναληπτική ανάπτυξη ή μάθηση και αυτή η μάθηση βασίζεται σε εμπειρικά δεδομένα. Για αυτόν τον λόγο δεν χρησιμοποιείται η λέξη «τεχνητή». Ένας άλλος λόγος για την αποφυγή της χρήσης «τεχνητού» είναι ότι υποδηλώνει προσομοίωση νοημοσύνης. Και αυτή είναι άλλη μία κοινή παρεξήγηση, αφού στόχος της δεν είναι να προσομοιώνει την νοημοσύνη. Ο πραγματικός στόχος είναι η κατανόηση πραγματικών ευφυών συστημάτων συνθέτοντάς τα. Όπως για παράδειγμα, μία προσομοίωση ενός σεισμού δεν είναι σεισμός, έτσι κι αυτό που απαιτείται είναι η δημιουργία πραγματικής νοημοσύνης, όπως για παράδειγμα η δημιουργία ενός πραγματικού σεισμού.

ΚΕΦΑΛΑΙΟ 3

ΗΘΙΚΗ

3.1. Εισαγωγή στην έννοια της ηθικής.

Ο όρος Ηθική, που υιοθετήθηκε και από τους ξένους, είναι ελληνικός και παράγεται από το ήθος, που σήμαινε αρχικά τη διαμονή ή το ενδιαίτημα (κατοικία) ανθρώπων και ζώων. Αργότερα δε, απέβαλε την αρχική του σημασία και προσέλαβε ειδική έννοια, που σημαίνει το εσωτερικό φρόνημα και τον χαρακτήρα, δηλαδή το μόνιμο και σταθερό τρόπο του πράττειν των ελεύθερων λογικών όντων. Ο ελληνικός αυτός όρος φαίνεται να παρήχθη από τη λέξη έθος (συνήθεια). Ήδη ο Αριστοτέλης δίδασκε, ότι οι αρετές βρίσκονται σε άμεση εξάρτηση από τη συνήθεια και την άσκηση. Έκτοτε ο όρος Ηθική επικράτησε στη φιλοσοφία και σήμαινε το σύνολο της περί τα ηθικά ζητήματα πραγματείας.²⁶

Η κατανόηση των ηθικών ζητημάτων απαιτεί την κατανόηση της ίδιας της ηθικής και ιδιαίτερα την λειτουργία της. Όμως τι πραγματικά σημαίνει ηθική και γιατί την χρειαζόμαστε; Μπορεί να συνδεθεί η ηθική με την υπολογιστική νοημοσύνη; Η ηθική ως κλάδος της φιλοσοφίας είναι "η φιλοσοφική σκέψη του ανθρώπου για τη θεωρητική θεμελίωση των πράξεών του τόσο προς τον εαυτό του, όσο και προς τους άλλους".²⁷ Είναι ένα "σύστημα αξιών και αρχών της πράξης που ισχύει για όλα τα ανθρώπινα όντα και για όλα τα όντα που έχουν τη δυνατότητα να κάνουν σχέδια και να παίρνουν αποφάσεις, με συνέπειες για όλες τις οντότητες που αισθάνονται πόνο".²⁸

Οι περισσότερες κοινωνικές και φιλοσοφικές έρευνες σε αυτόν τον τομέα υποστηρίζουν ότι ο κύριος ή ακόμα και μοναδικός σκοπός της ηθικής είναι να διατηρηθεί μια κοινωνική ομοιογένεια σε λειτουργικό βαθμό για την προστασία του κοινωνικού συνόλου. Φιλόσοφοι και κοινωνιολόγοι, όπως ο Εμίλ Ντιρκχάιμ, υποστηρίζουν ότι χωρίς την ηθική ως μηχανισμό της ανθρώπινης κοινωνικής συνοχής, μια κοινωνία αποσυντίθεται. Η κοινωνία δομείται στη βάση της κοινωνικής αλληλεγγύης μέσα από την ηθική συνένωση των ανθρώπων.²⁹

Η ηθική μπορεί να οριστεί από διαφορετικές πλευρές. Τη νομική πλευρά που προστατεύει τις αξίες με νόμους του κράτους, τη θρησκευτική πλευρά που αναφέρεται στις αρχές της θρησκείας, τη κοινωνική και πολιτιστική πλευρά που αναφέρεται στις αξίες που χαρακτηρίζουν μια κοινωνία, και την επαγγελματική πλευρά που στηρίζεται σε αξίες που διέπουν την επαγγελματική δραστηριότητα.

Μία έρευνα της Phi Delta Kappa το 2001³⁰, αναφέρει ως σημαντικότερες αξίες, που πρέπει να καλλιεργούνται από τους γονείς για να διαπλάσουν τα παιδιά τους ηθικούς χαρακτήρες, την τιμιότητα, την ικανότητα συνεργασίας, την υπευθυνότητα, την ανοχή, την αυτοπειθαρχία και τις δημοκρατικές αξίες. Σε γενικές γραμμές υπάρχουν κι άλλες αξίες που θεωρούνται αναγκαίες για τη διαμόρφωση ηθικής, όπως σεβασμός, το αίσθημα δικαίου, ο διαχωρισμός του καλού από το κακό μέσα από την εμπειρία και την συναναστροφή με

²⁶ Ι.Καρδάση, «Η Ηθική Στην Αρχαία Ελλάδα», Θησαυρός Γνώσεων Και Ευσεβείας(Ι), Οκτώβριος 2014

²⁷ Παπανούτσου Ε.Π., ΗΘΙΚΗ, Τόμος Ι, 5η έκδοση, Εκδόσεις Δωδώνη, 1995

²⁸ Δραγώνα-Μονάχου Μυρτώ, «Σύγχρονη Ηθική Φιλοσοφία-Ο Αγγλόφωνος Στοχασμός», Εκδόσεις Ελληνικά Γράμματα, 1995

²⁹ Γάγγας Σ., «Ο Εμίλ Ντιρκχάιμ και η Συγκρότηση της Ηθικής στην Νεωτερικότητα», σε Σ.Μ. Κονιόρδο, Αθήνα Εκδόσεις Gutenberg, 2010

³⁰ Brynildssen S, «Character education through children's literature», Αύγουστος 2002

άλλους ανθρώπους, η ισότητα, η κριτική σκέψη και η ηθική κρίση η οποία είναι απαραίτητη για μια ηθική ζωή.

Υπάρχουν βέβαια και συγκρουόμενες αξίες που δυσκολεύουν τις ηθικές κρίσεις, όπως οι υποχρεώσεις, οι προσωπικές δεσμεύσεις, τα συμφέροντα. Αυτές οι συγκρούσεις είναι εμφανείς σε πολλούς τομείς της καθημερινότητάς μας και ανάλογα τη σπουδαιότητα που έχει ένα θέμα, δίνεται βαρύτητα και σε διαφορετικές αξίες. Αυτό που είναι σίγουρο είναι ότι οι περίπλοκες ηθικές καταστάσεις που αντιμετωπίζουμε κατά διαστήματα, χρειάζονται κάτι παραπάνω από απλή γνώση των ηθικών κανόνων και αξιών. Ένα κατασκεύασμα υπολογιστικής νοημοσύνης, θα μπορούσε άραγε να σταθεί επάξια δίπλα στην ανθρώπινη ηθική κρίση;

3.2. Ιστορική αναδρομή στην έννοια της ηθικής.

Οι πρώτοι που ασχολήθηκαν σοβαρά με το θέμα της ηθικής ήταν ο Σωκράτης, ο Πλάτωνας και οι έλληνες σοφιστές. Αργότερα, ο Αριστοτέλης ήταν αυτός που της έδωσε το όνομα "ηθική". Σαν κλάδος όμως της φιλοσοφίας, καθιερώνεται από τους Στωικούς. Οι ηθικοί κανόνες που διαμορφώνουν την ορθή συμπεριφορά, διαφέρουν από κοινωνία σε κοινωνία, από εποχή σε εποχή. Οι αρχαίοι Έλληνες, για παράδειγμα, πίστευαν ότι η δουλεία ήταν αποδεκτή. Πλέον όμως σκεφτόμαστε διαφορετικά. Τον δέκατο ένατο αιώνα υπήρχε η διαφωνία αν πρέπει οι γυναίκες και οι μαύροι να ψηφίζουν. Τον εικοστό όμως αιώνα, δεν υπήρχαν τέτοια ηθικά διλήματα. Με το ίδιο σκεπτικό, υπάρχει πιθανότητα, οι άνθρωποι της σημερινής εποχής να θεωρηθούμε από τους πολιτισμούς του μέλλοντος, ως ηθικά ατελή όντα.³¹

Οι Στωικοί διαίρεσαν τη φιλοσοφία σε τρεις διαφορετικούς κλάδους: τη λογική, τη φυσική και την ηθική. Αυτή η διαίρεση των Στωικών επικράτησε μέχρι την Αναγέννηση αλλά και μετά. Την αποδέχθηκε και ο μεγάλος Γερμανός φιλόσοφος Immanuel Kant, ο φιλόσοφος που τόσο επηρέασε τη διαμόρφωση της σύγχρονης ηθικής των Δυτικών Κοινωνιών. Ο Kant υποστήριξε ότι υπάρχουν δύο επίπεδα πραγματικότητας, το επίπεδο των φαινομένων που αντιστοιχεί στην επιστήμη και το επίπεδο των νοουμένων που αντιστοιχεί στην ηθική. Το φαινόμενο επίπεδο δημιουργείται από το ανθρώπινο πνεύμα. Το νοούμενο επίπεδο υπερβαίνει την ανθρώπινη διάνοια και αντιστοιχεί σε μια πνευματική πραγματικότητα, που στηρίζει την ηθική και θρησκευτική ζωή του ανθρώπου.³²

Φαίνεται ότι οι διαφορετικές ανάγκες, και τα διαφορετικά συμφέροντα των ανθρώπων, όπως και οι συνθήκες που επικρατούν σε κάθε κοινωνία, καθορίζουν τους κώδικες ηθικής συμπεριφοράς. Έτσι, οι σοφιστές θεωρούσαν ότι είναι αδύνατο να υπάρξει ένας κοινός ηθικός κώδικας για όλες τις κοινωνίες. Αυτή την άποψη όμως, αμφισβήτησε ο Σωκράτης. Πρέσβευε ότι η αρετή ταυτίζεται με την γνώση και την αλήθεια. Αφού η αλήθεια είναι μία και μπορούν όλοι να τη γνωρίσουν, τότε όλοι θα συμφωνήσουν με τους ίδιους ηθικούς κανόνες. Πίστευε ότι η άγνοια είναι η αιτία του κακού κι όχι το προσωπικό όφελος.

Ο Kant πάλι, υποστήριξε ότι η αρετή δεν πηγάζει ούτε από το χαρακτήρα, ούτε από τον αυτοέλεγχο, ούτε από το Θεό, αλλά από την καλή βούληση του ανθρώπου, η οποία είναι το ύψιστο αγαθό και προϋπόθεση για όλα τα υπόλοιπα αγαθά.³³ Ο Nick Bostrom και ο Eliezer Yudkowsky υποστηρίζουν ότι την ηθική, την καθορίζουν συστήματα αξιών και κανόνων που είναι γενικώς αποδεκτά, είτε γιατί θεωρούνται αυτονόητα, είτε γιατί φαίνονται

³¹ Nick Bostrom, Eliezer Yudkowsky, 2011

³² Δ.Β.ΡΑΖΗΣ Δ/ντής Ογκολογικής Κλινικής Νοσοκομείου "ΥΓΕΙΑ", «Ηθική και Βιολογική Εξέλιξη», 1996

³³ Immanuel Kant, «Θεμέλια της μεταφυσικής των Ηθών», 1785

σωστά και δίκαια. Το σύνολο αυτών των κανόνων, επηρεάζει σημαντικά το χαρακτήρα, τις επιλογές, αλλά και τη συμπεριφορά των ανθρώπων.³⁴

3.3. Ηθική της Νοημοσύνης

Αναμφίβολα, η ηθική σχετίζεται άμεσα με τη νοημοσύνη, ή τουλάχιστον την προαπαιτεί ως αναγκαία συνθήκη, αφού σίγουρα δεν περιμένουμε ούτε απαιτούμε ηθική στάση από αυτούς που δεν έχουν αναπτύξει νοημοσύνη, όπως τα ζώα και τα μωρά παιδιά, ούτε από αυτούς που την έχουν χάσει.

«Ηθική νοημοσύνη είναι η ικανότητα να ξεχωρίζεις το καλό απ' το κακό, αλλά και να σέβεσαι τις αξίες των άλλων». Είναι η ικανότητα που έχει κάποιος, να παίρνει αποφάσεις που είναι σωστές και ωφέλιμες για τον εαυτό του, αλλά και για τον κόσμο γύρω του. «Είναι η αρετή που μας μαθαίνει να κάνουμε όλα αυτά που οφείλουμε και όχι αυτά που μας αρέσει να κάνουμε». Μην κάνεις αυτό που δε θέλεις να σου κάνουν, λέει κι ο Χριστιανισμός. Ηθική νοημοσύνη είναι εκείνο το «είδος της εξυπνάδας που συνδυάζει τη δύναμη του χαρακτήρα με την ορθή κρίση».

Ο καθηγητής και συγγραφέας Dr Michele Borba στο βιβλίο που έγραψε «Building Moral Intelligence (Χτίζοντας την Ηθική Νοημοσύνη)», αναφέρει τους τρόπους που οι γονείς μπορούν να διδάξουν τα τέκνα τους τις «αρετές της ηθικής νοημοσύνης». Αυτές οι αρετές είναι οι παρακάτω επτά:

- Ενσυναίσθηση: «Είναι η ικανότητα να μπορούμε να αναγνωρίσουμε τα συναισθήματα και τις ανάγκες των άλλων».
- Συνείδηση: «Είναι η ικανότητα να αναγνωρίζεις το σωστό και να μπορείς να το κάνεις».
- Αυτοέλεγχος: «Είναι η ικανότητα να ελέγχουμε τις σκέψεις και τις πράξεις μας και να λειτουργούμε με τον τρόπο που έχουμε μάθει πως είναι ο σωστός».
- Σεβασμός: «Είναι το να δείχνεις πως εκτιμάς τους άλλους φροντίζοντάς τους, αλλά και μιλώντας τους με πολιτισμένο τρόπο».
- Καλοσύνη: «Το να νοιάζεσαι για την ευημερία και τα αισθήματα των άλλων».
- Ανεκτικότητα: «Το να σέβεσαι τις αξίες και τα δικαιώματα όλων των ανθρώπων ακόμη κι αν διαφωνείς με τα πιστεύω τους».
- Δικαιοσύνη: «Το να διαλέγεις την ανοιχτόμυαλη συμπεριφορά και να συμπεριφέρεσαι δίκαια στους άλλους».

Μπορεί να υφίστανται διάφοροι τύποι νοημοσύνης όμως η ηθική, αν και παραμένει λιγότερο γνωστή από όλους, δεν παύει να είναι το ίδιο σημαντική. Κι αυτό γιατί είναι κομμάτι της ζωής μας, της ίδιας μας της καθημερινότητας. Είναι η ικανότητα να σέβεσαι, να εκτιμάς, αλλά και να δείχνεις συναισθήματα. Όσοι τη διαθέτουν, ξεχωρίζουν για τη δίκαιη και ορθή τους κρίση, αλλά και τη δύναμη και σταθερότητα του χαρακτήρα τους. Η ηθική νοημοσύνη είναι η νοημοσύνη που: «συνδυάζει τη δύναμη του χαρακτήρα με την ορθή κρίση».

Το χτίσιμο της ηθικής νοημοσύνης προσφέρει στους ανθρώπους ευημερία και ενισχύει την ελπίδα. Είναι ίσως η πιο ανθρωποκεντρική νοημοσύνη και συμβάλει στο χτίσιμο μια δίκαιης, ευτυχισμένης και γεμάτης νόημα κοινωνίας. Αυτοί που έχουν εκπαιδευτεί ηθικά, είναι και αυτοί που μπορούν να πολεμήσουν την βία, να καταπολεμήσουν τον αυθορμητισμό, την δυστυχία και να δώσουν υπόσταση σε μια άδεια ζωή. Άραγε μια μηχανική νοημοσύνη θα μπορέσει ποτέ να την αποκτήσει;

³⁴ Nick Bostrom, Eliezer Yudkowsky, 2011

3.4. Ηθική της Τεχνητής Νοημοσύνης

Η ηθική της τεχνητής νοημοσύνης είναι το τμήμα της ηθικής της τεχνολογίας που αφορά τα ρομπότ και άλλα τεχνητά νοήμονα όντα. Είναι τυπικά διαιρεμένο σε ρομποτική ηθική, δηλαδή στην μέριμνα για την ηθική συμπεριφορά των ανθρώπων όταν σχεδιάζουν, κατασκευάζουν, χρησιμοποιούν και διαχειρίζονται τεχνητά νοήμονα όντα και σε ηθική της μηχανής, η οποία ασχολείται με την ηθική συμπεριφορά τεχνητών ηθικών παραγόντων (AMAs). Τα συστήματα τεχνητής νοημοσύνης είναι κατασκευασμένα από ανθρώπους για να εκπληρώσουν ορισμένους στόχους. Απαιτούνται θεωρίες, μέθοδοι και αλγόριθμοι για την ενσωμάτωση ηθικών, κοινωνικών και νομικών και αξιών στις τεχνολογικές εξελίξεις στην τεχνητή νοημοσύνη, σε όλα τα στάδια ανάπτυξης (ανάλυση, σχεδιασμός, κατασκευή, ανάπτυξη και αξιολόγηση). Αυτά τα ηθικά, κοινωνικά και νομικά κριτήρια σημαίνουν πλαίσια μέσω των οποίων κατευθύνονται οι επιλογές σχεδιασμού και ρύθμισης των συστημάτων τεχνητής νοημοσύνης, όπως και κανόνες εμπλοκής με τον ανθρώπινο παράγοντα.³⁵

Ενώ είναι γενικά αποδεκτό ότι τα σημερινά συστήματα τεχνητής νοημοσύνης δεν έχουν ηθική υπόσταση, δεν είναι σαφές τι ακριβώς θεμελιώνει το ηθικό. Συνήθως προτείνονται δύο κριτήρια που συνδέονται σημαντικά με την ηθική κατάσταση, είτε χωριστά είτε σε συνδυασμό: η επίγνωση και η γνώση (μάθηση) ή σοφία που καθορίζει τον ανθρώπινο παράγοντα. Η επίγνωση είναι η ικανότητα για φαινομενική εμπειρία (qualia), όπως και η ικανότητα να αισθάνεται κανείς πόνο και να υποφέρει. Η σοφία, ή μάθηση αντιπροσωπεύει ένα σύνολο δυνατοτήτων που συνδέονται με υψηλότερη ευφυΐα, όπως η αυτογνωσία και η ύπαρξη ενός παράγοντα που ανταποκρίνεται στον λόγο³⁶.

Η ηθική της τεχνητής νοημοσύνης ασχολείται με τις ηθικές επιπτώσεις των ευφύων μηχανών. Τα ηθικά ζητήματα που συνδέονται με το ΑΙ πολλαπλασιάζονται και βρίσκονται συνεχώς στην επικαιρότητα. Οι ευφείς μηχανές εμφανίζονται στην καθημερινότητά μας, σαν εργαλεία για την διερεύνηση της συνείδησης και άλλες πτυχές της γνωστικής λειτουργίας που συμβάλλουν στην ηθική υπόσταση. Αυτό έχει βαθιές επιπτώσεις στην κατανόηση της ηθικής, και πόσο μάλλον της ηθικής των συστημάτων στο βαθμό που αυτά τα συστήματα συμφωνούν με τα πειραματικά αποτελέσματα. Επιπλέον, τίθεται το ζήτημα του υπεύθυνου ή και αξιοκατάκριτου συστήματος τεχνητής νοημοσύνης που λειτουργεί ανοιχτά στην κοινωνία. Αυτό έχει και πάλι βαθιές επιπτώσεις για τα συστήματα της ηθικής που πρέπει να “φιλοξενήσουν” την ηθική της τεχνητής νοημοσύνης.

Σίγουρα για να είναι κάποιος ηθικός δεν απαιτείται να έχει μόνο την λογική και την διάθεση να ακολουθήσει τον ηθικό κανόνα κάθε στιγμή, αλλά να έχει και την απαραίτητη διαίσθηση για να αναγνωρίζει τι θα μπορούσε να έχει ηθική σημασία, έτσι ώστε να επιστρατέψει τους ηθικούς κανόνες και να χειριστεί την κατάσταση ανάλογα. Απαιτείται, δηλαδή, «ηθική ευαισθησία» που είναι χαρακτηριστικό λογικών όντων που είναι ταυτόχρονα και βιολογικοί οργανισμοί.

³⁵ Dignum, Virginia. «Ethics in artificial intelligence: introduction to the special issue». *Ethics and Information Technology (Springer Netherlands)* **20** (1): 1–3. doi:10.1007/s10676-018-9450-z. ISSN 1572-8439.

³⁶ Bostrom, Nick. Yudkowsky, Eliezer (2014). «The ethics of artificial intelligence». Στο: Keith Frankish and William M. Ramsey, επιμ. *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, σελ. 322–323. ISBN 978-0-521-69191-8

Συνεπώς η ηθική ευαισθησία δεν μπορεί να καταγραφεί επακριβώς ώστε να εφαρμοστεί από οποιοδήποτε τεχνητά νοήμον σύστημα, αλλά ούτε θα μπορούσε να προκύψει μέσω της κατάλληλης εκπαίδευσης σε κάποιο νευρωνικό δίκτυο, καθώς η ηθική ευαισθησία εμπλέκει την ηθική κατανόηση. Για παράδειγμα, θεωρείται ηθικό να λέμε πάντα την αλήθεια και επίσης έχουμε καθήκον να προστατεύουμε τους φίλους μας, τι πρέπει να κάνουμε όμως σε περίπτωση που τα δύο αυτά καθήκοντα έρθουν σε σύγκρουση; Αν κάποιος παρανοϊκός μ' ένα τσεκούρι στο χέρι με ρωτούσε πού βρίσκεται ο φίλος μου, η πρώτη μου ενστικτώδης αντίδραση θα ήταν να του πω ψέματα. Λέγοντας την αλήθεια θα αμελούσα το καθήκον που έχω να προστατέψω το φίλο μου. Από την άλλη μεριά όμως, το να πω ψέματα, ακόμα και σε μια τέτοια ακραία κατάσταση θεωρείται ανήθικη πράξη. Πώς λοιπόν θα έπρεπε ένα τεχνητά νοήμων σύστημα να προγραμματιστεί ώστε να είναι ηθικό από όλες τις απόψεις;

3.4.1. Έννοια της ρομποτικής.

Από την αρχαιότητα έχει απασχολήσει η συζήτηση για τη δεοντολογία σε σχέση με την αντιμετώπιση των μη ανθρώπινων και μη έμβιων όντων και την δυνατότητά τους στην «πνευματικότητα». Με την ανάπτυξη των μηχανημάτων και τελικά των ρομπότ, αυτή η φιλοσοφία εφαρμόστηκε επίσης στην ρομποτική. Αν και ο όρος "roboethics" δημιουργήθηκε από τον Gianmarco Veruggio, ρομποτικός επιστήμονας, πρωτοπόρος της Διαδικτυακής Ρομποτικής, το 2002, ο όρος αναπτύχθηκε από τον Isaac Asimov στο πλαίσιο των επιστημονικής φαντασίας έργων του.

Η "Ρομπο-ηθική" (Roboethics) είναι μια σύντομη έκφραση για την ηθική της ρομποτικής. Ασχολείται με την συμπεριφορά των ανθρώπων σε σχέση με τα ρομπότ. Δηλαδή με το πώς οι άνθρωποι σχεδιάζουν, κατασκευάζουν, χρησιμοποιούν και συμπεριφέρονται στα ρομπότ και σε άλλα τεχνητά νοήμονα όντα. Ειδικότερα, οι κύριοι τομείς που εμπλέκονται στην Ρομποτική είναι: η ρομποτική, η επιστήμη των υπολογιστών, η τεχνητή νοημοσύνη, η φιλοσοφία, η ηθική, η θεολογία, η βιολογία, η φυσιολογία, οι γνωστικές επιστήμες, οι νευροεπιστήμες, το δίκαιο, η κοινωνιολογία, η ψυχολογία και το βιομηχανικό σχέδιο. Ο σχεδιασμός της Ρομποτικής απαιτεί τη συνεργασία των εμπειρογνομόνων των παραπάνω ειδικοτήτων, και αυτών που εργάζονται σε επιτροπές οι οποίες πρέπει να προσαρμόσουν τους νόμους και τους κανονισμούς για τα προβλήματα που προκύπτουν από τα επιστημονικά και τεχνολογικά επιτεύγματα σε Ρομποτική και Τεχνητή Νοημοσύνη (Artificial Intelligence-AI).

Ο Peter M. Asaro, φιλόσοφος επιστημών, τεχνολογίας και μέσων μαζικής ενημέρωσης, ασχολείται με την τεχνική νοημοσύνη και ρομποτική ως μια μορφή ψηφιακών μέσων, καθώς και με τους τρόπους που η τεχνολογία μεσολαβεί στις κοινωνικές σχέσεις και διαμορφώνει την εμπειρία μας για τον κόσμο. Όσον αφορά τον τομέα της ηθικής αναφέρει: *"Υπάρχουν τουλάχιστον τρία πράγματα που θα μπορούσαμε να εννοούμε με τον όρο "ηθική στη ρομποτική: τα ηθικά συστήματα που ενσωματώνονται στα ρομπότ, την ηθική των ανθρώπων που σχεδιάζουν και χρησιμοποιούν τα ρομπότ και την ηθική του πώς οι άνθρωποι αντιμετωπίζουν τα ρομπότ. Για την καλύτερη προσέγγιση στην ηθική των ρομπότ θα πρέπει να εξετάσουμε και τους τρεις τομείς, κάτι που προϋποθέτει την αντιμετώπιση του ρομπότ ως ένα κοινωνικο-τεχνικό σύστημα. Με τον τρόπο αυτό, είναι δυνατό να σκεφτούμε ότι το ρομπότ μπορεί να κινηθεί μεταξύ της μη συναίσθησης της ηθικής και της πλήρως αυτόνομης ηθικής και να εξελίσσεται, όσο αποκτά μεγαλύτερες δυνατότητες και ηθική περιπλοκότητα. Πολλά από τα ζητήματα που αφορούν την κατανομή των ευθυνών σε σύνθετα κοινωνικο-τεχνικά συστήματα μπορούν να αντιμετωπιστούν καλύτερα εξετάζοντάς τα περισσότερο από νομική άποψη, παρά από ηθική. Αυτό επειδή το συνολικό ενδιαφέρον μας για την ηθική των ρομπότ πρέπει να είναι*

η πρακτική της πρόληψης ενός ρομπότ να κάνει κακό, καθώς και η πρόληψη των ανθρώπων για την αποφυγή άδικων ευθυνών για τις πράξεις του".³⁷

Επίσης εκφράζει κάποια σοβαρά ερωτήματα που στο άμεσο μέλλον θα πρέπει να αντιμετωπίσουμε και να απαντήσουμε: "Σκεφτείτε την περίπτωση που σε ένα ρομπότ δοθούν δύο αντικρουόμενες εντολές από δύο διαφορετικούς ανθρώπους. Ποιον θα πρέπει να υπακούσει; Τον ιδιοκτήτη; Το πιο ασφαλές κοινωνικά άτομο; Το πιο ηθικό αίτημα; Ή το άτομο που του αρέσει καλύτερα; Ή μήπως θα έπρεπε να ακολουθήσει το αίτημα που εξυπηρετεί τα δικά του συμφέροντα καλύτερα; Και όσον αφορά εμάς, θα μας πείραζε το ότι πρόκειται να λάβει την δική του απόφαση;"

Οι άνθρωποι αντιμετωπίζουν τέτοια διλήμματα όλη την ώρα. Η πρακτική ηθική βοηθάει στην επιχειρηματική δραστηριότητα της παροχής μέσων για την επίλυση αυτών των ζητημάτων. Υπάρχουν διάφορα συστήματα για τη διαμόρφωση αυτών των ηθικών μελετών, αλλά, τελικά, εξαρτάται από το άτομο ως προς το ποιο σύστημα, θα χρησιμοποιήσει. Η διαφορά με τα ρομπότ είναι ότι τα ηθικά συστήματα πρέπει να ενσωματωθούν, μετά από επιλογή από τους σχεδιαστές, σε κάθε τεχνολογικό σύστημα, τα οποία είναι χτισμένα συστήματα, όπως το "dilemmas". Ακόμη όμως και σε συστήματα που θα μπορούσαν να μάθουν ηθικούς κανόνες ή συμπεριφορές, δεν είναι σαφές ότι θα μπορούν να χαρακτηριστούν ως αυτόνομοι ηθικοί παράγοντες, και η συσκευή που υπογράφει αυτές τις μεθόδους μάθησης θα εξακολουθούσε να είναι υπεύθυνη για την αποτελεσματικότητά τους. Ίσως κάποια μέρα ένα ρομπότ να μπορέσει να φτάσει σε ένα σημείο ανάπτυξης, ώστε οι σχεδιαστές του και οι προγραμματιστές να μην είναι πλέον υπεύθυνοι για τις πράξεις του, όπως και οι γονείς ενός παιδιού που δεν θεωρούνται υπεύθυνοι για τις πράξεις του όταν αυτό έχει ενηλικιωθεί.

Αυτό είναι σίγουρα μια ενδιαφέρουσα δυνατότητα, τόσο επειδή θέτει το ζήτημα του τι θα κάνει ένα ρομπότ με μια αυτόνομη ηθική, όσο και το ερώτημα πώς ένας τέτοιος παράγοντας μπορεί να μοιάζει. Έχουν υπάρξει ζωντανόι λογοτεχνικοί και φιλοσοφικοί λόγοι σχετικά με τα όρια για κατηγορίες όπως "ζωντανό ή μη ζωντανό" και "συνειδητό ή μη συνειδητό", και αυτά φαίνεται να είναι στενά συνδεδεμένα με την ηθική του ρομπότ. Δεν είναι όμως σαφές ότι μια ικανοποιητική κατανομή αυτών των ορίων θα απλοποιήσει τα ηθικά ζητήματα. Πράγματι, η ηθική μπορεί να περιπλέξει. Ενώ θα μπορούσε να αποδειχθεί ότι είναι δυνατόν να δημιουργηθούν πραγματικά αυτόνομοι τεχνητοί ηθικοί παράγοντες, αυτό φαίνεται να είναι θεωρητικά και τεχνολογικά μια πρόκληση για το προβλέψιμο μέλλον.³⁸

3.4.2. Ηθική της μηχανής (machine ethics)

Η "ηθική της μηχανής" (machine ethics), ασχολείται με την συμπεριφορά των ίδιων των ρομπότ, είτε θεωρούνται τεχνητοί ηθικοί παράγοντες είτε όχι και αν τελικά έχουν δικαιώματα. Καθώς τα ρομπότ έχουν γίνει πιο προηγμένα και εξελιγμένα, εμπειρογνώμονες και ακαδημαϊκοί έχουν να διερευνήσουν ερωτήματα όπως τι μπορεί να ισχύει για την ηθική συμπεριφορά των ρομπότ και αν τα ρομπότ θα μπορούσαν να είναι σε θέση να διεκδικήσουν κάθε είδους κοινωνικά, πολιτιστικά, ηθικά ή νομικά δικαιώματα.

Όλος ο τεχνολογικός κόσμος μιλάει για το ότι η τεχνητή νοημοσύνη καλπάζει και ότι στα επόμενα χρόνια τα ρομπότ θα κατακλύσουν τα νοσοκομεία μας ως αδελφές, τα καταστήματα ως πωλητές, τις επιχειρήσεις ως γραμματείς και τα σπίτια μας ως οικιακοί βοηθοί. Αλλά πώς μπορείς να εμπιστευθείς τη διάδραση με ανθρώπους σε μηχανές χωρίς

³⁷ Peter M. Asaro, *International Review of Information Ethics* 6 (12):9-16), «What should we want from a robot ethic» Peter M. Asaro, 2006

³⁸ *International Review of Information Ethics Vol. 6*, «What should we want from a robot ethic», IRIE© by IRIE – all rights reserved www.i-r-i-e.net 10, ISSN 1614-1687 Peter M. Asaro, Δεκέμβριος 2006

αισθήματα και ηθική; Πώς θα μπορούν οι «έξυπνες μηχανές» να αντιμετωπίζουν τις ποικίλες διαφορές συμπεριφοράς ή ψυχοσωματικής κατάστασης των ανθρώπων αν δεν έχουν την απαραίτητη «συναισθηματική νοημοσύνη»;

Η πρώτη εργασία πάνω στο θέμα της ηθικής της μηχανής, είχε τίτλο «Χρησιμοποιώντας Ιδεατή Πραγματικότητα για την Αξιολόγηση Ηθικών Αποφάσεων σε Σενάρια Οδικής Κυκλοφορίας», διεξήχθη από ερευνητές του γερμανικού πανεπιστημίου Osnabrück και δημοσιεύθηκε στο *Frontiers in Behavioral Neuroscience*.³⁹ Κύριο θέμα της ήταν τα ρομποτικά αυτοκίνητα χωρίς οδηγό. Αυτά τα αυτο-οδηγούμενα οχήματα πρωτοεμφανίστηκαν τον Οκτώβριο του 2015, από την εταιρεία Tesla του Elon Musk. Έκτοτε είδαμε πολλούς ανταγωνιστές της να δοκιμάζουν τέτοια οχήματα σε αυτοκινητοδρόμους αρχικά και σε δρόμους του αστικού τοπίου αργότερα.

Οι προβλέψεις θέλουν τα, χωρίς ανθρώπινο οδηγό, αυτοκίνητα να κυριαρχούν στις πόλεις του 2040. Αλλά, τότε, πώς θα παίρνουν τις σωστές αποφάσεις σε ό,τι αναπάντεχο μπορεί να προκύψει; Για παράδειγμα, όπως διαπίστωσε πρόσφατα η Volvo, το υποπρόγραμμα ανίχνευσης μεγάλων ζώων που έχει στο πειραματικό αυτο-οδηγούμενο μοντέλο της στάθηκε ανίκανο να προϋπολογίσει την κίνηση των καγκουρό στην Αυστραλία, γιατί δεν είχε ξαναδεί ζωντανά που προχωρούσαν πηδώντας. Το ερώτημα εδώ είναι βάσει ποιών ηθικών κριτηρίων θα παίρνουν οι μηχανές την απόφαση όταν αντιμετωπίζουν αναπόφευκτα διλήμματα;

Επειδή οι προγραμματιστές των αυτο-οδηγούμενων οχημάτων κατέλειξαν στο συμπέρασμα ότι «δεν είναι δυνατόν να προγραμματίσεις την ηθική», το γερμανικό υπουργείο Μεταφορών και Ψηφιακής Υποδομής (BMVI) αποδέχθηκε ότι «η ηθική ανθρώπινη συμπεριφορά δεν μπορεί να μοντελοποιηθεί» και συνέταξε έναν κατάλογο 20 ηθικών αρχών βάσει των οποίων θα κρίνονται τα όποια αναπότρεπτα ατυχήματα τέτοιων οχημάτων κατά της ανθρώπινης ζωής.

Ο κύριος λόγος που οι προγραμματιστές τεχνητής νοημοσύνης είχαν καταλήξει σε αδυναμία μοντελοποίησης της ηθικής είχε να κάνει με το ότι δεν τους δίνονταν «σταθερές εκτίμησης». Δηλαδή, στα ατυχήματα κατά ζωής κρίνονται τα πάντα κατά περίπτωση, βάσει πλαισίου συνθηκών, με άλλοτε τη στεγνή λογική να κυριαρχεί και άλλοτε τον συναισθηματισμό των ατόμων. Η συγκεκριμένη έρευνα του Πανεπιστημίου του Οσναμπρουκ (Osnabrück) θέλησε να διαπιστώσει αν όντως η ηθική των ατυχημάτων εξαρτάται αναπόδραστα από τις συνθήκες και, άρα, είναι όντως μη μοντελοποιήσιμη. Φόρεσαν λοιπόν κράνη ιδεατής πραγματικότητας (Virtual Reality) σε 105 εθελοντές ηλικίας 18-60 ετών και τους εξέθεσαν σε προσομοιώσεις ποικιλίας οδηγικών διλημμάτων. Κατά τα λεγόμενα από τον επικεφαλής της έρευνας, τον Leon Sütfeld, «μέχρι τώρα υποθέταμε ότι οι ηθικές αποφάσεις εξαρτώνται κατά πολύ από το πλαίσιο των συνθηκών και συνεπώς ότι δεν μπορούν να μοντελοποιηθούν ή να περιγραφούν αλγοριθμικά. Αλλά βρήκαμε ακριβώς το αντίθετο. Η συμπεριφορά των ανθρώπων σε διλημματικές καταστάσεις μπορεί να μοντελοποιηθεί βάσει ενός σχετικά απλού μοντέλου "αξίας της ζωής" που είχε ο κάθε συμμετέχων για κάθε άνθρωπο, ζώο ή άψυχο αντικείμενο (που εμπλεκόταν στο υποτιθέμενο ατύχημα)». Αυτό σημαίνει ότι η ανθρώπινη ηθική συμπεριφορά μπορεί κάλλιστα να αποδοθεί με αλγόριθμους, που επίσης κάλλιστα μπορούν να γίνουν προγράμματα για τα ρομποτικά αυτοκίνητα.

Οι συγγραφείς της μελέτης επισήμαναν ότι τα αυτο-οδηγούμενα οχήματα είναι μόνο η αρχή και προειδοποίησαν ότι χαράζει μια νέα εποχή με ανάγκη για σαφείς κανόνες. Διαφορετικά, οι μηχανές θα αρχίσουν να παίρνουν αποφάσεις χωρίς εμάς. Χαρακτηριστικά, ένας άλλος κύριος ερευνητής της ομάδας, ο καθηγητής Gordon Pira, εξέφρασε τον εξής προβληματισμό: «Από τη στιγμή που φαίνεται ότι είναι δυνατόν οι μηχανές να προγραμματιστούν για να παίρνουν ηθικές αποφάσεις, είναι ζωτικής σημασίας να ξεκινήσει η

³⁹ βλ. journal.frontiersin.org/article/10.3389/fnbeh.2017.00122/full

κοινωνία μια επείγουσα και σοβαρή συζήτηση. Πρέπει να αναρωτηθούμε κατ' αρχάς αν οι μηχανές θα πρέπει να αποκτήσουν τη δυνατότητα ηθικής αξιολόγησης. Εάν ναι, τότε πώς θα συμπεριφέρονται; Μιμούμενες τις ανθρώπινες αποφάσεις ή υπακούοντας σε θεωρίες ηθικής;».

Παρόμοιο προβληματισμό αλλά για κάτι πολύ πιο προχωρημένο εξέφρασαν οι ερευνητές της δεύτερης ομάδας που συνέγραψε την εργασία «Βοήθεια, ελπίδα και υπερβολή: ηθικές διαστάσεις της νευροπροσθετικής». Τη διεξήγαγαν ερευνητές των πανεπιστημίων Freiburg και Tübingen (Γερμανία), Washington (ΗΠΑ), Wyss Geneva (Ελβετία), Keio (Ιαπωνία) και Ottawa (Καναδάς) και τη δημοσίευσαν στο περιοδικό *Science*.⁴⁰ Το δικό τους πεδίο εφαρμογής ήταν οι συσκευές εγκεφαλικής διεπαφής (Brain-Machine Interface, BMI) και τα εγκεφαλικά εμφυτεύματα μέσω των οποίων οι επιχειρήσεις τεχνολογίας προσβλέπουν ότι μελλοντικά οι πελάτες τους θα «διατάσσουν τις μηχανές».

Όπως δήλωσε ο επικεφαλής της έρευνας και διευθυντής του Κέντρου Βιολογίας και Νευρομηχανικής Wyss της Γενεύης, καθηγητής John Donoghue, «παρότι ακόμη δεν κατανοούμε πλήρως τον τρόπο με τον οποίο λειτουργεί ο εγκέφαλος, βρισκόμαστε όλο και πιο κοντά στη δυνατότητα να αποκωδικοποιούμε αξιόπιστα ορισμένα εγκεφαλικά σήματα. Δεν μας επιτρέπεται να εφησυχάσουμε για το τι μπορεί να σημάνει αυτό για την κοινωνία. Πρέπει να εξετάσουμε προσεκτικά τις συνέπειες της συμβίωσης με ημι-έξυπνες μηχανές ελεγχόμενες από τον εγκέφαλό μας και θα πρέπει να είμαστε έτοιμοι να διασφαλίσουμε την ασφαλή και ηθική χρήση τους με κάποιους μηχανισμούς».⁴¹

3.5. Ανάλυση της έννοιας: υπολογιστική ηθική (CE:computer ethics).

Η πληροφορική (υπολογιστική) ηθική ως αντικείμενο μελέτης και προβληματισμού πρωτοεμφανίζεται στα τέλη της δεκαετίας του σαράντα και συγκεκριμένα στο έργο του καθηγητή Norbert Wiener, ο οποίος προέβλεψε τις κοινωνικές και ηθικές συνέπειες της πληροφορικής τεχνολογίας. Στα βιβλία του *Cybernetics: or control and communication in the animal and the machine* (1948) και *The Human Use of Human Beings* (1950) θεμελιώνει τις θεωρητικές βάσεις της πληροφορικής ηθικής. Ο Wiener προβλέπει διορατικά ότι η επανάσταση της πληροφορικής θα έχει τεράστιες επιπτώσεις στην κοινωνία, γι' αυτό και την ορίζει ως "δεύτερη βιομηχανική επανάσταση". Μία δεκαετία αργότερα ο Donn ξεκίνησε την μελέτη της μη ηθικής χρήσης των ηλεκτρονικών υπολογιστών και συγκέντρωσε πολλές περιπτώσεις ηλεκτρονικού εγκλήματος (Parker 1968, 1979, 1990).

Σύμφωνα με επιστήμονες, η πληροφορική ηθική ανήκει στην εφαρμοσμένη ηθική, το παρακλάδι εκείνο της κανονιστικής ηθικής το οποίο μελετά πρακτικά και καθημερινά προβλήματα όπως αυτά της περιβαλλοντικής ηθικής και της βιοηθικής. Η πληροφορική ηθική μπορεί να οριστεί ως η συστηματική μελέτη των ηθικών προβλημάτων που προκύπτουν από την είσοδο της πληροφορικής τεχνολογίας στην κοινωνία.⁴²

Ο όρος "πληροφορική ηθική" (computer ethics) χρησιμοποιήθηκε για πρώτη φορά από τον καθηγητή Walter Maner ο οποίος ασχολήθηκε με τη φιλοσοφική μελέτη εκείνων των ηθικών προβλημάτων που δημιουργούνται, μεταλλάσσονται και επιβαρύνονται από την πληροφορική τεχνολογία. Ο Maner στο βιβλίο του *Starter Kit in Computer Ethics*⁴³, μελετά την παιδαγωγική σημασία της πληροφορικής ηθικής. Η πρώτη ολοκληρωμένη μελέτη παρουσιάζεται το 1985 στο άρθρο του James Moore "What is Computer Ethics?". Στο άρθρο

⁴⁰ βλ. science.sciencemag.org/content/356/6345/1338.full

⁴¹ Καφαντάρης Τάσος, «Η ηθική της μηχανής», Ιούλιος 2017

⁴² Βικιπαίδεια, Απρίλιος 2016

⁴³ Walter Maner, 1978

του αυτό ο Moore, προσφέρει έναν πρώτο ορισμό της πληροφορικής ηθικής και παρουσιάζει μια ολοκληρωμένη διάσταση του φιλοσοφικού της αντικειμένου. Ορίζει λοιπόν την πληροφορική ηθική, ως *"την ανάλυση της φύσης και της κοινωνικής επιρροής της πληροφορικής τεχνολογίας και την αντίστοιχη διαμόρφωση και δικαιολόγηση των κανόνων της προς ηθική χρήση αυτής της τεχνολογίας"*.

Η ηθική είναι ένα σύνολο ηθικών αρχών που κυβερνούν τη συμπεριφορά μιας ομάδας ή ενός ατόμου. Επομένως, η ηθική των υπολογιστών είναι οι ηθικές αρχές που ρυθμίζουν τη χρήση των υπολογιστών. Μερικά κοινά ζητήματα της ηθικής υπολογιστών περιλαμβάνουν τα δικαιώματα πνευματικής ιδιοκτησίας, όπως το ηλεκτρονικό περιεχόμενο, θέματα που έχουν να κάνουν με τη μυστικότητα και την ασφάλεια των προσωπικών δεδομένων, και πώς οι υπολογιστές έχουν επιπτώσεις στην κοινωνία, ανάλογα με τη χρήση τους.

Παραδείγματος χάριν, ενώ είναι εύκολο κάποιος να αναπαραγάγει το ηλεκτρονικό (ή ψηφιακό) περιεχόμενο, η ηθική υπολογιστών θα πρότεινε ότι είναι λανθασμένο να συμβεί χωρίς την έγκριση του συντάκτη. Και ενώ μπορεί να είναι δυνατό για ένα συγκρότημα ηλεκτρονικών υπολογιστών, να προσεγγίσει τις προσωπικές πληροφορίες κάποιου, η ηθική υπολογιστών θα συμβούλευε ότι μια τέτοια δράση είναι ανήθικη. Με την προϋπόθεση ότι αναπτύσσεται η τεχνολογία και τα υπολογιστικά συστήματα συνεχίζουν να ασκούν μεγάλη επίδραση στην κοινωνία, η ηθική των υπολογιστών προωθεί τη συζήτηση για το πόση επίδραση θα έπρεπε οι υπολογιστές να έχουν σε τομείς όπως η τεχνητή νοημοσύνη και η ανθρώπινη επικοινωνία.

Δεδομένου ότι ο κόσμος των υπολογιστών εξελίσσεται, η ηθική των υπολογιστών συνεχίζει να δημιουργεί ηθικά πρότυπα για την αντιμετώπιση των νέων ζητημάτων που προκύπτουν από τις νέες τεχνολογίες: *"Η ουσιαστική δυσκολία για τη φιλοσοφική θέση της ηθικής των υπολογιστών, είναι ένα μεθοδολογικό πρόβλημα που οι τυποποιημένες ηθικές θεωρίες δεν μπορούν εύκολα να προσαρμοστούν για να εξετάσουν και το CE απαιτεί ένα εννοιολογικό θεμέλιο ως ηθική θεωρία. Τι είναι καλό για μια οντότητα πληροφοριών; Αυτή είναι η ηθική ερώτηση που υποβάλλεται από το φιλοσοφικό θεμελιώδες αντίστοιχο του CE. Η απάντηση δίνεται από μια μινιμαλιστική θεωρία: Η ηθική των πληροφοριών υποστηρίζει ότι υπάρχει κάτι πιο στοιχειώδες και θεμελιώδες από τη ζωή και τον πόνο, δηλαδή η ύπαρξη, που γίνεται κατανοητή ως πληροφορία, και την εντροπία, και ότι οποιαδήποτε οντότητα πληροφοριών πρόκειται να αναγνωριστεί ως κέντρο μιας ελάχιστης ηθικής αξίωσης, η οποία αξίζει την αναγνώριση και πρέπει να βοηθήσει να ρυθμιστεί η εφαρμογή οποιασδήποτε διαδικασίας πληροφοριών που την περιλαμβάνει. Η ηθική των πληροφοριών (IE: Information Ethics), που σχετίζεται με τα πεδία της ηθικής των υπολογιστών και την φιλοσοφία της πληροφορίας, μπορεί να παρέχει μια πολύτιμη προοπτική από την οποία να προσεγγίσει, με τη διορατικότητα και την επαρκή διάκριση, όχι μόνο τα ηθικά προβλήματα στο CE, αλλά και ολόκληρη τη σειρά των εννοιολογικών και ηθικών φαινομένων που διαμορφώνουν την ηθική ομιλία".*⁴⁴

⁴⁴ Luciano Floridi, "Ethics and Information Technology" Publisher Kluwer Academic Publishers Hingham, MA, USA 1998-01-01) <http://dl.acm.org/citation.cfm?id=594590>

ΚΕΦΑΛΑΙΟ 4

ΣΥΝΔΕΣΗ ΗΘΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ

4.1. Η ηθική υπόσταση των υπολογιστικών συστημάτων

Η διαχωριστική γραμμή μεταξύ δημιουργού και δημιουργήματος έχει γίνει θολή τον τελευταίο καιρό, χάρη στις συνεχείς εξελίξεις που οδηγούν σε νέες γενιές ρομπότ. Οι "μηχανές" πλέον, είναι αρκετά έξυπνες για να εφεύρουν δικές τους τεχνολογίες, να σχεδιάσουν τους δικούς τους αλγόριθμους καλύτερα, πιο γρήγορα και πιο έξυπνα. Αυτά είναι τα ρομπότ του μέλλοντος που θα μοιραστούν τον κόσμο μαζί μας. Σύμφωνα με την έρευνα που διενεργήθηκε από την Υπηρεσία του "UK Office of Science and Innovation's Horizon Scanning Centre", το ρομπότ θα μπορούσε μια μέρα να διεκδικήσει τα δικαιώματα ενός πολίτη, και τελικά την αλλαγή των νόμων.⁴⁵

Μια επιστημονική ομάδα δήλωσε ότι περίπου το 2019, είναι πιθανό να δημιουργηθεί ένας εγκέφαλος ρομπότ. Ο κοινωνικός αντίκτυπος των ευφών ρομπότ είναι επίσης το θέμα που πραγματεύτηκε το 2010 το ντοκιμαντέρ «Plug & Pray». Ο Vernor Vinge καθηγητής Μαθηματικών και Επιστήμης Υπολογιστών, ανέφερε ότι μπορεί να υπάρξει μια εποχή όπου οι υπολογιστές και τα ρομπότ θα είναι εξυπνότερα από τους ανθρώπους, το οποίο αποκαλεί ως "η Μοναδικότητα" (the Singularity). Ο ίδιος πίστευε ότι αυτό μπορεί ενδεχομένως να είναι πολύ επικίνδυνο για τον άνθρωπο. Το συγκεκριμένο θέμα εξετάζεται από μια φιλοσοφία που ονομάζεται Singularitarianism. Σημειωτέον ορισμένα ρομπότ έχουν αποκτήσει διάφορες μορφές ημι-αυτονομίας, συμπεριλαμβανομένης της ικανότητάς τους να βρίσκουν τις πηγές τροφοδοσίας από μόνα τους και να είναι σε θέση να επιλέγουν ανεξάρτητα στόχους για να επιτεθούν με όπλα. Επίσης ορισμένοι ιοί υπολογιστών μπορούν να αποφύγουν την εξάλειψη, το οποίο αναφέρεται ως "νοημοσύνη κατσαρίδας" (cockroach intelligence).⁴⁶

Η δυνατότητα δημιουργίας μηχανών σκέψης, εγείρει μια σειρά από ηθικά ζητήματα. Τα ερωτήματα αυτά αφορούν, αφενός αν διασφαλίζεται ότι οι μηχανές αυτές δεν βλάπτουν τους ανθρώπους και τα άλλα ηθικά όντα, και αφετέρου την ηθική υπόσταση των ίδιων των μηχανών. Είναι ευρέως αποδεκτό ότι τα σημερινά συστήματα υπολογιστικής νοημοσύνης, δεν έχουν καμία ηθική υπόσταση. Μπορούμε δηλαδή να αλλάξουμε, να αντιγράψουμε, να διαγράψουμε, ή να χρησιμοποιήσουμε τα προγράμματα ηλεκτρονικών υπολογιστών που μας αρέσουν. Τι ακριβώς όμως, προσδίδει ηθική υπόσταση; Τα δύο κριτήρια που συνήθως προτείνονται ως σημαντικά, και συνδέονται με την ηθική υπόσταση, είτε χωριστά είτε σε συνδυασμό, είναι η ευαισθησία και η σοφία. Θα προσπαθήσουμε να εξηγήσουμε παρακάτω, τι εννοούμε με αυτούς τους όρους.

Ως ευαισθησία ορίζεται η ικανότητα που αποκτούμε από εμπειρία, όπως για παράδειγμα η αίσθηση πόνου. Ενώ ως σοφία ορίζεται ένα σύνολο ικανοτήτων που συνδέονται με την υψηλότερη νοημοσύνη, όπως για παράδειγμα η αυτογνωσία. Με λίγα λόγια, η σοφία αναφέρεται στα χαρακτηριστικά της προσωπικότητας του ατόμου. Μια κοινή άποψη είναι ότι πολλά ζώα έχουν ευαισθησία και ως εκ τούτου έχουν κάποια ηθική υπόσταση, αλλά ότι μόνο τα ανθρώπινα όντα έχουν τη νοημοσύνη, η οποία τους δίνει μια υψηλότερη ηθική υπόσταση. Σ' αυτή την περίπτωση όμως, θα πρέπει να αντιμετωπίσουμε την ύπαρξη κάποιων εξεζητημένων περιπτώσεων, όπως για παράδειγμα τα ανθρώπινα βρέφη ή οι άνθρωποι με σοβαρή νοητική καθυστέρηση. Επειδή έχουν χαμηλή νοημοσύνη, πρέπει να

⁴⁵ Easton, 1 Ιουνίου 2015, «Roboethics – We need Unoversal Robots Rights, Ethics and Legislation»

⁴⁶ Αθανασάκη, 2015

θεωρήσουμε άραγε ότι δεν πληρούν τα κριτήρια για την σοφία; Κι αν πάλι θεωρήσουμε ότι έχουν προσωπικότητα, γιατί να μη θεωρήσουμε ότι για τον ίδιο λόγο έχουν και τα ζώα;

Αν δεχτούμε αυτό το σκεπτικό, ένα σύστημα ΑΙ θα έχει κάποια ηθική υπόσταση αν έχει την ικανότητα ευαισθησίας, όπως την ικανότητα να αισθάνεται πόνο. Ένα αισθανόμενο υπολογιστικό σύστημα, ακόμη και αν δεν έχει μητρική γλώσσα και υψηλή μόρφωση, μοιάζει περισσότερο με ένα ζωντανό ζώο παρά με ένα άψυχο παιχνίδι. Είναι λάθος να προκαλούμε πόνο σε ένα ποντίκι, αν δεν υπάρχουν ισχυροί ηθικά επιτακτικοί λόγοι για να το πράξουμε. Αν διευρύνουμε αυτό το σκεπτικό, τότε το ίδιο θα πρέπει να ισχύει για κάθε αισθανόμενο υπολογιστικό σύστημα. Με βάση τα παραπάνω, εάν εκτός από ευαισθησία, ένα σύστημα έχει επίσης νοημοσύνη παρόμοια με εκείνη ενός κανονικού ενήλικα ανθρώπου, τότε ποιός μπορεί να αρνηθεί ότι θα έχει πλήρη ηθική υπόσταση, ισοδύναμη με εκείνη των ανθρώπινων όντων;

Κάποιοι υποστηρίζουν ότι αν δύο όντα έχουν την ίδια λειτουργικότητα και την ίδια συνειδητή εμπειρία, και διαφέρουν μόνο ως προς το πώς ήρθαν στην ύπαρξη και στον τρόπο κατασκευής τους, τότε έχουν την ίδια ηθική υπόσταση. Σύμφωνα με αυτούς, το αντίθετο θα ισοδυναμούσε με μια θέση παρόμοια με το ρατσισμό όπως για παράδειγμα η διάκριση για το χρώμα του δέρματος. Με άλλα λόγια, υποστηρίζουν, ότι δεν υπάρχει καμία ηθική διαφορά ανάμεσα σε ένα ον που είναι κατασκευασμένο από πυρίτιο ή άνθρακα και ο εγκέφαλός του χρησιμοποιεί ημι-αγωγούς ή νευροδιαβιβαστές, και σε ένα ανθρώπινο όν, αρκεί να υπάρχουν τα δυο παραπάνω κοινά χαρακτηριστικά, η ίδια λειτουργικότητα και η ίδια συνειδητή εμπειρία.

Ακόμη και εκείνοι που είναι αντίθετοι στην αναπαραγωγική κλωνοποίηση του ανθρώπου για ηθικούς ή θρησκευτικούς λόγους, γενικά αποδέχονται ότι, ένας ανθρώπινος κλώνος θα έχει την ίδια ηθική υπόσταση όπως και κάθε άλλο ανθρώπινο βρέφος. Αν γεννηθεί ένα ανθρώπινο έμβρυο χωρίς εγκέφαλο από καθαρά φυσικά αίτια, θεωρείται ότι έχει ηθική υπόσταση. Αν γεννηθεί ένα ανεγκέφαλο παιδί μετά από μία τεχνητή παρέμβαση πρέπει άραγε να έχει την ίδια ηθική υπόσταση;

Η Αρχή της Οντογένεσης των Διακρίσεων ισχυρίζεται ότι ο τρόπος που δημιουργήθηκε ένα ον δεν έχει ουσιαστική σχέση με την ηθική υπόστασή του. Δεν αμφισβητεί όμως ότι ο δημιουργός έχει υποχρεώσεις και καθήκοντα ως προς το εν λόγω ον. Όπως οι γονείς έχουν συγκεκριμένα ηθικά καθήκοντα για το παιδί τους, και όχι για άλλα παιδιά ακόμα κι αν μοιάζουν με το δικό τους, ομοίως, οι δημιουργοί ή οι ιδιοκτήτες ενός συστήματος ΑΙ με ηθική υπόσταση, μπορεί να έχουν ηθικά καθήκοντα στο τεχνητό μυαλό του συστήματος αυτού, χωρίς βέβαια να έχουν σε ένα άλλο τεχνητό μυαλό, ακόμα και αν είναι ποιοτικά όμοιο.

Εάν γίνουν γενικά αποδεκτές αυτές οι αντιλήψεις, τότε θα πρέπει να αλλάξει τελείως ο τρόπος που αντιμετωπίζουμε τα τεχνητά μυαλά, εφαρμόζοντας τις ίδιες ηθικές αρχές που χρησιμοποιούμε για να προσδιορίσουμε τα καθήκοντά μας στο ανθρώπινο περιβάλλον. Στο βαθμό που τα ηθικά καθήκοντα πηγάζουν από ζητήματα ηθικής υπόστασης, οφείλουμε να αντιμετωπίζουμε ένα τεχνητό μυαλό ακριβώς με τον ίδιο τρόπο που αντιμετωπίζουμε ένα ποιοτικά όμοιο φυσικό ανθρώπινο μυαλό σε μια παρόμοια κατάσταση.

Ακόμα και αν δεχτούμε τη στάση αυτή όμως, πρέπει να αντιμετωπίσουμε μια σειρά από νέα ηθικά ερωτήματα τα οποία αφήνουν αναπάντητα οι προαναφερθείσες αντιλήψεις. Νέα ηθικά ζητήματα προκύπτουν επειδή τα τεχνητά μυαλά μπορούν να έχουν πολύ διαφορετικές ιδιότητες από τα συνήθη μυαλά των ανθρώπων ή των ζώων. Πρέπει να εξετάσουμε πώς αυτές οι καινούριες ιδιότητες θα μπορούσαν να επηρεάσουν την ηθική υπόσταση του τεχνητού μυαλού και τι θα σήμαινε να σεβόμαστε την ηθική υπόσταση τέτοιων μαγικών μυαλών.

Αν θέλουμε να δούμε σοβαρά την ανάπτυξη προηγμένης τεχνητής νοημοσύνης, θα πρέπει να αντιμετωπίσουμε αυτή την πρόκληση. Αν τα μηχανήματα στο μέλλον καταλήξουν να γίνουν πιο δυνατά, πιο γρήγορα, πιο αξιόπιστα και πιο εξυπνότερα από τους ανθρώπους,

τότε η ηθική πειθαρχία της μηχανής πρέπει να δεσμευτεί να αναζητήσει τα ανώτερα ανθρώπινα αισθήματα και όχι μόνο τις ανθρώπινες ισοδυναμίες. Η προοπτική της υπολογιστικής νοημοσύνης με υπεράνθρωπη ευφυΐα και υπερφυσικές ικανότητες, μας φέρνει αντιμέτωπους με την αναγκαιότητα να κατασκευαστεί ένας αλγόριθμος που θα εξάγει υπερ-ηθική συμπεριφορά. Οι προκλήσεις αυτές μπορεί να φαίνονται σαν σενάρια επιστημονικής φαντασίας, αλλά φαίνεται αναμενόμενο ότι θα τις συναντήσουμε στο άμεσο μέλλον.

Ο δρ Δημήτριος Καλλές, πρόεδρος της Ελληνικής Εταιρείας Τεχνητής Νοημοσύνης και επίκουρος καθηγητής στο Ανοιχτό Πανεπιστήμιο, πρεσβεύει ότι: *«Αυτό που θα λείπει πάντοτε από τις μηχανές είναι η συνείδηση, δηλαδή η ικανότητα που έχει ο άνθρωπος να κατανοεί τον εαυτό του. Εξάλλου, για να κωδικοποιήσουμε σε αλγόριθμους την ανθρώπινη νοημοσύνη, θα έπρεπε να ξέρουμε ακριβώς πώς αυτή λειτουργεί, κάτι που νομίζω πως είναι αδύνατον να γίνει»*. Ο δρ Ιωάννης Βλαχάβας, καθηγητής Πληροφορικής στο ΑΠΘ, συμφωνεί πως η συνείδηση θα είναι πάντα αυτό που θα ξεχωρίζει τον άνθρωπο από τα δημιουργήματά του: *«Ακόμη κι αν ένα λογισμικό εξελιχθεί ώστε να απαντά στις πιο περίπλοκες ερωτήσεις, ποτέ δεν θα αποκτήσει αυτογνωσία, δηλαδή ποτέ δεν "θα ξέρει πως ξέρει, όπως ο άνθρωπος. Επίσης, η πρωτοβουλία θα παραμείνει πάντοτε ένα ανθρώπινο χαρακτηριστικό, με την έννοια ότι ποτέ δεν θα καταφέρει μια μηχανή να κάνει κάτι διαφορετικό απ' ό,τι έχει προγραμματιστεί»*.⁴⁷

4.2 Συνείδηση και Υπολογιστική Νοημοσύνη.

Οι θεωρίες που προσπαθούν να περιγράψουν το τι ακριβώς είναι η συνείδηση είναι πολλές και διάφορες. Σύμφωνα με τον Robert Lawrence Kuhn υπάρχουν 5 Εναλλακτικές αιτίες συνείδησης.

1. Η συνείδηση είναι εντελώς φυσική, αποκλειστικά το προϊόν του φυσικού εγκεφάλου, το οποίο στα βαθύτερα του επίπεδα περιλαμβάνει τα πεδία και τα σωματίδια της θεμελιώδους φυσικής. Αυτός είναι ο «φυσικισμός» ή ο «υλισμός» και είναι συντριπτικά η επικρατούσα θεωρία των επιστημόνων.
2. Η συνείδηση είναι ένα ανεξάρτητο, μη αναγωγικό χαρακτηριστικό της φυσικής πραγματικότητας που υπάρχει εκτός από τα πεδία και τα σωματίδια της θεμελιώδους φυσικής. Αυτό μπορεί να πάρει τη μορφή μιας νέας, ανεξάρτητης (πέμπτης) φυσικής δύναμης ή μιας ριζικά νέας οργάνωσης της πραγματικότητας.
3. Η συνείδηση είναι ένα μη αναγωγικό χαρακτηριστικό κάθε φυσικού πεδίου και σωματιδίου της θεμελιώδους φυσικής. Ό, τι υπάρχει ένα είδος «πρωτο-συνείδησης», το οποίο, σε ορισμένα συσσωματώματα και κάτω από ορισμένες συνθήκες, μπορεί να δημιουργήσει εσωτερική συνειδητοποίηση σε ανθρώπινο επίπεδο. Αυτός είναι ο «πανψυχισμός», μια από τις παλαιότερες θεωρίες στη φιλοσοφία του νου (που επιστρέφει στις προ-σύγχρονες κινούμενες θρησκείες και στους αρχαίους Έλληνες). Ο πανψυχισμός, με διάφορες μορφές, είναι μια ιδέα που αναβιώνουν ορισμένοι σύγχρονοι φιλόσοφοι ως απάντηση στο φαινομενικά ανυπόστατο "σκληρό πρόβλημα" της συνείδησης.
4. Η συνείδηση απαιτεί μια ριζικά ξεχωριστή, μη φυσική ουσία που είναι ανεξάρτητη από έναν φυσικό εγκέφαλο, έτσι ώστε η πραγματικότητα να αποτελείται από δύο ριζικά

⁴⁷ Κώστας Δελγιάννης, «COMPUTERS», Δεκέμβριος 2014

διαφορετικά μέρη - φυσικές και μη φυσικές ουσίες, διαιρέσεις, διαστάσεις ή επίπεδα ύπαρξης. Αυτό είναι "δυϊσμός". Ενώ η ανθρώπινη συνείδηση απαιτεί και έναν φυσικό εγκέφαλο και αυτή τη μη φυσική ουσία (με κάποιο τρόπο να συνεργάζεται), μετά από το θάνατο του σώματος και τη διάλυση του εγκεφάλου, αυτή η μη φυσική ουσία θα μπορούσε να διατηρήσει κάποιο είδος συνειδητής ύπαρξης. (Αν και αυτή η μη φυσική ουσία παραδοσιακά ονομάζεται «ψυχή» - όρος που έχει έντονες θεολογικές συνέπειες - μια ψυχή δεν είναι το μοναδικό είδος που μπορεί να είναι μια τέτοια μη φυσική ουσία).

5. Η συνείδηση είναι η απόλυτη πραγματικότητα. Το μόνο πράγμα που είναι πραγματικό. Όλα τα υπόλοιπα, συμπεριλαμβανομένου ολόκληρου του φυσικού κόσμου, προέρχονται από μια συνολική "κοσμική συνείδηση". Κάθε μεμονωμένη περίπτωση συνείδησης - ανθρώπινου, ζωικού, ρομποτικού ή άλλου - αποτελεί μέρος αυτής της κοσμικής συνείδησης. Οι ανατολικές θρησκείες, γενικά, υιοθετούν αυτήν την άποψη.⁴⁸

Ο ίδιος παραθέτει εάν αυτές οι υπερδιάνοιες θα μπορούσαν να συνειδητοποιήσουν και να έχουν εσωτερική συνείδηση σε κάθε μία από τις αιτίες που προαναφέρθηκαν. Ο Robert Lawrence Kuhn λοιπόν ισχυρίζεται ότι:

1. Εάν η συνείδηση είναι απολύτως φυσική, τότε θα ήταν σχεδόν σίγουρο ότι οι μη βιολογικές ευφυΐες με υπερβολική ΑΙ θα είχαν το ίδιο είδος εσωτερικής συνειδητοποίησης που έχουμε. Επιπλέον, καθώς το ΑΙ θα φτάσει την «Singularity-Μοναδικότητα» και θα γίνει αναποτελεσματικά πιο περίπλοκο από τον ανθρώπινο εγκέφαλο, θα μπορούσε πιθανότατα να εκφράσει μορφές συνείδησης υψηλότερες από αυτές που μπορούσαμε σήμερα να φανταστούμε.
2. Εάν η συνείδηση είναι ένα ανεξάρτητο, μη αναγωγικό χαρακτηριστικό της φυσικής πραγματικότητας, τότε θα παραμείνει ένα ανοιχτό ερώτημα εάν οι μη βιολογικές ευφυΐες θα μπορούσαν ποτέ να βιώσουν την αληθινή εσωτερική συνείδηση. (Θα εξαρτηθεί από τη βαθιά φύση του χαρακτηριστικού που προκαλεί τη συνειδητότητα και αν αυτό το χαρακτηριστικό θα μπορούσε να χειριστεί με την τεχνολογία).
3. Αν η συνείδηση είναι μια μη αναγωγική ιδιότητα κάθε στοιχειώδους φυσικού πεδίου και σωματιδίου (πανψυχισμός), τότε φαίνεται πιθανό ότι οι μη βιολογικές ευφυΐες με υπερβολικό ΑΙ θα μπορούσαν να βιώσουν την αληθινή εσωτερική συνείδηση (επειδή η συνείδηση θα ήταν ένα εγγενές μέρος του ιστού της φυσικής πραγματικότητας).
4. Εάν η συνείδηση είναι μια ριζικά ξεχωριστή, μη φυσική ουσία που δεν καθορίζεται από τον φυσικό κόσμο (δυϊσμός), τότε θα ήταν αδύνατο το υπερβολικό ΑΙ (μόνο), ανεξάρτητα από το πόσο εξελιγμένο είναι, να μπορούσε να βιώσει ποτέ την αληθινή εσωτερική συνειδητοποίηση.
5. Αν η συνείδηση είναι τελική πραγματικότητα (κοσμική συνείδηση), τότε οτιδήποτε μπορεί να είναι συνειδητό, συμπεριλαμβανομένων των μη βιολογικών ευφυιών.

Σύμφωνα με το φιλόσοφο, John Searle, τα προγράμματα ηλεκτρονικών υπολογιστών δεν θα μπορούσαν ποτέ να έχουν μυαλό ή να συνειδητοποιήσουν με την ανθρώπινη έννοια, ακόμη και αν δημιουργούν ισοδύναμες συμπεριφορές και αλληλεπιδράσεις με τον εξωτερικό κόσμο. Ο Warren Brown, ψυχολόγος στο Fuller Theological Seminary και μέλος του

⁴⁸ βλ. Deepak Chopra για σύγχρονα επιχειρήματα ότι η τελική πραγματικότητα είναι η συνείδηση.

Ινστιτούτου Ερευνών Εγκεφάλου της UCLA, στήριξε την «ενσωματωμένη γνώση, την ενσωματωμένη συνείδηση», στο ότι «η βιολογία είναι το πλουσιότερο υπόστρωμα για την ενσυναίσθηση της συνείδησης». Αλλά δεν αποκλείει ότι η συνείδηση «μπορεί να ενσωματωθεί σε κάτι μη βιολογικό». Από την άλλη πλευρά, ο Μπράουν στήριξε ότι «η συνείδηση μπορεί να είναι ένα ιδιαίτερο είδος οργάνωσης του κόσμου που απλά δεν μπορεί να αναπαραχθεί σε ένα μη βιολογικό σύστημα».

Μία άλλη εκδοχή ανέθεσε ο νευροεπιστήμονας Christof Koch, πρόεδρος και επικεφαλής επιστημονικός υπεύθυνος του ινστιτούτου Allen για την εγκεφαλική επιστήμη, ισχυρίζοντας: "Εφόσον μπορούμε να αναπαράγουμε τις ίδιες «σχετικές σχέσεις» μεταξύ όλων των σχετικών νευρώνων στον εγκέφαλο, πιστεύω ότι θα μπορούμε να αναδημιουργήσουμε την συνείδηση. Το δύσκολο μέρος είναι, τι εννοούμε με τον όρο «σχετικές σχέσεις;». Πρέπει να αναπαραγάγουμε όλες αυτές τις σχέσεις με την καλωδίωση ("connectome") του εγκεφάλου σε ένα διαφορετικό μέσο. Αν μπορούμε να κάνουμε όλα αυτά στο σωστό επίπεδο, αυτή η κατασκευή λογισμικού θα ήταν συνειδητή".

Τέλος, αξίζει να σημειωθεί η άποψη του ερευνητή Ray Kurzweil, ο οποίος πιστεύει ότι «θα φτάσουμε σε ένα σημείο όπου οι υπολογιστές θα αποδείξουν την πλούσια σειρά των συναισθηματικά λεπτών συμπεριφορών που βλέπουμε στους ανθρώπους · θα είναι πολύ έξυπνοι και θα ισχυρίζονται ότι είναι συνειδητοί. θα ενεργήσουν με συνειδητό τρόπο, θα μιλήσουν για τη συνείδησή τους και θα το συζητήσουν μόνο με τον τρόπο που εσύ και εγώ κάνουμε και έτσι η φιλοσοφική συζήτηση θα είναι εάν ή όχι συνειδητά - και θα συμμετέχουν στη συζήτηση".⁴⁹

4.3 Αυτόνομα συστήματα και ηθική.

Στη χαρραγή της δημιουργίας της αυτόνομης τεχνητής νοημοσύνης ικανής να παράγει λογισμούς, ο Στέφεν Χόκινγκ προειδοποίησε για τους κινδύνους που υπάρχουν στο τέλος της διαδρομής. Όπως είχε τονίσει σε συνέντευξή του στο BBC, "η ανάπτυξη πλήρους τεχνητής νοημοσύνης θα μπορούσε να σημάνει το τέλος του ανθρώπινου γένους". Είπε ακόμα ότι: "Από τη στιγμή που οι άνθρωποι θα αναπτύξουν την τεχνητή νοημοσύνη, αυτή θα απογειωθεί από μόνη της και θα ανασχεδιάζεται σε ολοένα και αυξανόμενο ρυθμό. Οι άνθρωποι, οι οποίοι περιορίζονται από μια αργή βιολογική εξέλιξη, δεν θα μπορούν να την ανταγωνιστούν και τελικά αυτή θα τους υποκαταστήσει". Παρά τους προβληματισμούς, τα ηθικά διλήμματα και τις φωνές που κάνουν λόγο για τους κινδύνους που ελλοχεύουν στην ανάπτυξη της τεχνητής νοημοσύνης, όπως αυτή του Στίβεν Χόκινγκ, υπάρχουν πολλοί οραματιστές που βλέπουν λύσεις από την ανάπτυξή της. Η τεχνητή νοημοσύνη τελικά θα λύσει μερικά από τα μεγαλύτερα προβλήματα του κόσμου; Ήδη πολλοί άνθρωποι εργάζονται ώστε να υλοποιηθεί αυτό.

Η Google Brain, για παράδειγμα, εστιάζει στο να μπορέσει να κάνει τις μηχανές να αυτοεκπαιδούνται μέσα από την κατανόηση των ψηφιακών δεδομένων. Σύμφωνα με ανάρτηση του Andrew Ng, ενός εκ των συνιδρυτών της Google Brain, στον ιστότοπο Quora, στόχος των προσπαθειών του φιλόδοξου πρότζεκτ στο πεδίο της τεχνητής νοημοσύνης, είναι να φτάσουν οι μηχανές στο σημείο όπου θα είναι σε θέση να λαμβάνουν αποφάσεις, όπως οι άνθρωποι. Αυτό μπορεί να γίνει με την συνεργασία πολλών αλγορίθμων, ή αλλιώς νευρωνικών δικτύων, για την επεξεργασία και καταχώρηση εικόνων και κειμένων γρήγορα και αποτελεσματικά.

⁴⁹ Robert Lawrence Kuhn, «When Robots Colonize the Cosmos, Will They Be Conscious? (Op-Ed) », Οκτώβριος 2015

Όπως σημειώνεται από τους ερευνητές της εταιρείας τεχνητής νοημοσύνης "DeepMind", η οποία ανήκει στην Google, τεχνητές νοημοσύνες με ενισχυμένη μάθηση που αλληλεπιδρούν σε πραγματικό χρόνο με ένα πολύπλοκο περιβάλλον σαν τον πραγματικό κόσμο, είναι μάλλον απίθανο να λειτουργούν πάντα άψογα. Σε περίπτωση που μια τέτοια νοημοσύνη λειτουργεί σε πραγματικό χρόνο υπό ανθρώπινη επίβλεψη, σημειώνεται, ίσως να είναι απαραίτητο για τον άνθρωπο χειριστή να μπορεί να πατήσει ένα μ κουμπί ασφαλείας, για να αποτρέψει την πραγματοποίηση επιβλαβών ενεργειών και να οδηγήσει την τεχνητή νοημοσύνη σε μια πιο ασφαλή κατάσταση. Επειδή όμως η τεχνητή νοημοσύνη πιθανόν να επιδιώξει να αποφύγει κάποια τέτοια διακοπή, εξουδετερώνοντας ίσως το κουμπί ασφαλείας, εξετάζεται από τους συγκεκριμένους επιστήμονες η δημιουργία ενός τρόπου που θα διασφαλίζει ότι η τεχνητή νοημοσύνη δεν θα μάθει να αποφεύγει ή να εμποδίζει τη διακοπή λειτουργίας της από το περιβάλλον ή από έναν άνθρωπο χειριστή.

Βασικό κομμάτι του reinforcement learning είναι ότι ο υπολογιστής διαμορφώνει την «προσωπικότητά» του μέσω των ανταμοιβών που λαμβάνει για τις επιτυχίες του, οπότε «διαβάζει» το περιβάλλον του και σταδιακά μαθαίνει πώς μπορεί να επιτύχει τις περισσότερες ανταμοιβές. Ωστόσο, αυτό σημαίνει ότι δεν μπορεί να καταλάβει πάντα πότε κάποιες από τις πράξεις του είναι επιβλαβείς και για αυτό χρειάζεται ένας άνθρωπος να το επιβλέπει και να το καθοδηγεί, που θα έχει επίσης τη δυνατότητα να σταματήσει την τεχνητή νοημοσύνη αν χρειαστεί.

Η λύση που προτείνεται είναι η "επιλεκτική αμνησία": Όταν ενεργοποιείται η ασφάλεια, το σύστημα συνεχίζει να λειτουργεί θεωρώντας ότι δεν θα διακοπεί ξανά. Οπότε αντί να μαθαίνει από την «τιμωρία» του, ανατρέχει στο σύστημα ανταμοιβών για να βγάλει νέα συμπεράσματα όσον αφορά στο πώς αξιολογούνται από πλευράς προτεραιοτήτων οι εργασίες. Έτσι δεν μαθαίνει να βγάζει λάθος συμπεράσματα. Άλλο ένα μέτρο που προτείνεται είναι να δίνεται στον υπολογιστή η εντύπωση πως έκανε την αλλαγή μόνος του, χωρίς οδηγία από έξω.⁵⁰

Σύμφωνα με δημοσίευμα του CNN, ο 34-χρονος Quoc Le, κάτοχος διδακτορικού πτυχίου από το Πανεπιστήμιο του Στάνφορντ στην επιστήμη των υπολογιστών και πρωτοπόρος στην τεχνητή νοημοσύνη, που εργάζεται εδώ και 4,5 χρόνια στην Google Brain, αναφέρει ότι "Υπάρχουν πολύ λίγοι άνθρωποι στον κόσμο που πραγματικά αντιλαμβάνονται πως μπορούν να κάνουν τις μηχανές να μάθουν και να σκεφτούν".⁵¹ Στην πραγματικότητα, οι μηχανές μπορούν και να μάθουν και να σκευτούν μόνες τους.

Απόδειξη ύπαρξης αυτών των μηχανών, είναι το Promobot IR77, ένα ρομπότ που διαθέτει ιδιαίτερα υψηλή τεχνητή νοημοσύνη, καθώς είναι σε θέση να "θυμάται και να μαθαίνει". Οι επιστήμονες είπαν ότι είχαν προγραμματίσει ολόκληρη τη σειρά του απλά να αποφεύγει εμπόδια, ωστόσο το συγκεκριμένο ρομπότ αποδείχθηκε πιο έξυπνο από τα άλλα και κατάφερε να αποδράσει δύο φορές, λαχταρώντας την ελευθερία του. Η περιπέτειά του ωστόσο έληξε άδοξα, καθώς αφού προκάλεσε πανικό στους ανυποψίαστους, ξέμεινε από μπαταρία στη μέση ενός δρόμου στο Περθ μετά από 45 λεπτά. «Μετά τις δύο αποδράσεις, διορθώσαμε το σύστημα AI (τεχνητής νοημοσύνης) και τώρα σκεφτόμαστε να διαλύσουμε τελείως το συγκεκριμένο ρομπότ», ανέφερε ο επικεφαλής του προγράμματος, καθησυχάζοντας τους πολίτες. Το ρομπότ δραπέτης είχε σχεδιαστεί να αλληλεπιδρά με τους ανθρώπους σαν να ήταν ένας από αυτούς, απαντώντας κανονικά σε ερωτήσεις, ενώ θυμάται κάθε πρόσωπο που έχει συναντήσει!

⁵⁰ ΤΕΧΝΟΛΟΓΙΑ, 08 Ιουνίου 2016

⁵¹ Newsroom , CNN Greece, 24 Φεβρουαρίου 2016



Εικόνα 4:

Τρόμος: Ρομπότ με ιδιαίτερα υψηλή νοημοσύνη δραπέτευσε για δεύτερη φορά από εργαστήριο

Αξίζει τέλος να σημειωθεί πως αυτή δεν είναι η πρώτη φορά που ρομπότ «αντιρρησίας» ξεφεύγει. Πριν από λίγο καιρό, είχε σημειωθεί στην Αυστρία η πρώτη αυτοκτονία ρομπότ, το οποίο βαρέθηκε να καθαρίζει και αφού τέθηκε μόνο του σε λειτουργία, σύμφωνα με τους ισχυρισμούς της οικογένειας στην οποία πρόσφερε τις υπηρεσίες του, ψήθηκε «ζωντανό» πάνω στην εστία της κουζίνας. Προκλήθηκε μάλιστα πυρκαγιά, ενώ οι πυροσβέστες επιβεβαίωσαν πως εντόπισαν υπολείμματα μηχανής πάνω στα μάτια.⁵²

Καθώς τα μηχανήματα αποκτούν όλο και περισσότερη υπολογιστική νοημοσύνη και αυτονομία, ο Colin Allen (φιλόσοφος και ερευνητής, καθηγητής στο Bloomington, πανεπιστήμιο της Indiana), και ο Russel (καθηγητής της επιστήμης των υπολογιστών), συμφωνούν ότι θα απαιτούν ολοένα και πιο εξελιγμένες ηθικές ικανότητες. Κάτι το οποίο μας φέρνει αντιμέτωπους με το πρώτο κολοσσιαίο εμπόδιο: Δεν υπάρχει ένα καθολικό πρότυπο ανθρώπινης ηθικής. Η ηθική είναι με πολιτισμικές ιδιαιτερότητες, συνεχώς εξελισσόμενη και αιώνια υπό συζήτηση. Εάν τα ρομπότ πρόκειται να ζήσουν σύμφωνα με έναν ηθικό κώδικα, ποιος θα είναι αυτός; Από τί θα αποτελείται; Ποιος αποφασίζει; Επιπλέον, οι ειδικοί σε θέματα ρομποτικής, έχουν ακόμα ένα εμπόδιο: Πώς να θέσουν τα ανθρώπινα ήθη στο μυαλό μίας μηχανής.

Σύμφωνα με τον Allen, συγγραφέα του βιβλίου Ηθικές Μηχανές (Moral Machines), η πιο άμεση μέθοδος είναι να προγραμματιστούν στο λογισμικό του ρομπότ, ρητοί κανόνες για τη συμπεριφορά. Οι κανόνες θα μπορούσαν να είναι συγκεκριμένοι, όπως για παράδειγμα οι Δέκα Εντολές του CEI (Computer Ethics Institute), ή οι Τρεις Νόμοι της Ρομποτικής του Asimov, κανόνες τους οποίους θα αναλύσουμε σε επόμενο κεφάλαιο. Θα μπορούσαν βέβαια να είναι πιο θεωρητικοί, επηρεαζόμενοι από την κατηγορική προστακτική του Ιμμάνουελ

⁵² newsbomb.gr, Ιούνιος 2016

Καντ⁵³ που διατυπώνεται χωρίς όρους και παίρνει τη μορφή καθολικής απαίτησης του ηθικού νόμου, η οποία έχει γενικό, αντικειμενικό, απόλυτο και υποχρεωτικό χαρακτήρα.

Επίσης μπορούσαν να επηρεαστούν, από την ωφελμιστική ηθική, σύμφωνα με την οποία σκοπός των πράξεων μας πρέπει να είναι η μεγαλύτερη κατά το δυνατόν ωφέλεια για τον μεγαλύτερο κατά το δυνατόν αριθμό ατόμων. Κυριότεροι εκπρόσωποι αυτής της ηθικής είναι ο Τζέρεμι Μπένθαμ⁵⁴ και ο Τζον Στιούαρτ Μιλλ.⁵⁵

Άλλοι επιστήμονες θεωρούν ότι το σημαντικότερο είναι να δίνονται στο μηχανήμα κωδικοποιημένες κατευθυντήριες οδηγίες πάνω στις οποίες θα βασίσει την λήψη των αποφάσεών του. Σε αυτήν την περίπτωση, ο μηχανικός διατηρεί τον έλεγχο του τι πρέπει να γνωρίζει το ρομπότ και του τι όχι. Ο Allen όμως ισχυρίζεται ότι τα ρομπότ θα αντιμετωπίσουν ένα πολύ μεγάλο υπολογιστικό φορτίο κατά τη λήψη γρήγορων αποφάσεων στον πραγματικό κόσμο. Χρησιμοποιώντας ως παράδειγμα τον πρώτο κανόνα του Asimov (μην βλάψετε τους ανθρώπους), εξηγεί: "Για να υπολογίσει εάν μία δεδομένη δράση βλάπτει ή όχι πραγματικά έναν άνθρωπο απαιτεί να είναι σε θέση να υπολογίσει όλες τις συνέπειες της δράσης στο μακρινό μέλλον". Κάτι που κάνουμε εμείς οι άνθρωποι, σαν απόρροια ετών πρακτικής εμπειρίας και εκπαίδευσης. "Ο Αριστοτέλης είπε ότι ο τρόπος για να γίνεις καλός άνθρωπος είναι η πρακτική κάνοντας καλά πράγματα", λέει ο Allen, και αυτός μπορεί να είναι επίσης, ο καλύτερος τρόπος για να γίνει ένα ρομπότ πολύ καλό.

Το πεδίο της μηχανικής μάθησης έχει κάνει μεγάλα άλματα. Οι ειδικοί της ρομποτικής χρησιμοποίησαν την ενισχυτική μάθηση, "καλό ρομπότ - κακό ρομπότ", με επιτυχία. Επίσης ο Russel εφηύρε μια άλλη τεχνική που ονομάζεται αντίστροφη ενισχυτική μάθηση, η οποία πάει τα πράγματα ένα βήμα παραπέρα. Χρησιμοποιώντας την μέθοδο του Russel, ένα ρομπότ παρατηρεί την συμπεριφορά μίας άλλης οντότητας, ενός ανθρώπου ή ακόμα και ενός άλλου ρομπότ, και εκτός του να μιμηθεί απλώς τις ενέργειές του, προσπαθεί να καταλάβει ποιος είναι ο βασικός στόχος.

Σύμφωνα με τον Russel αυτή η προσέγγιση θα μπορούσε να επιτρέψει στα ρομπότ να μάθουν την ανθρώπινη ηθική μέσα από τα ανθρώπινα μέσα, όπως τις ταινίες, τα μυθιστορήματα, τα ειδησεογραφικά προγράμματα και τις τηλεοπτικές εκπομπές. Όλα αυτά αποτελούν έναν τεράστιο θησαυρό πληροφοριών σχετικά με τις ανθρώπινες αξίες. Με αυτόν τον τρόπο τα μηχανήματα μαθαίνουν όπως ένα παιδί. Έτσι θα μπορούσαν να μάθουν τί κάνει τους ανθρώπους χαρούμενους, τί τους κάνει λυπημένους, τί κάνουν και μπαίνουν φυλακή και τί κάνουν για να κερδίσουν μετάλλια. Σε λίγα χρόνια λοιπόν οι τηλεοράσεις θα πρέπει να παρέχουν φίλτρα και για ρομπότ, όχι μόνο για ενήλικες.

Βασικός στόχος του Russel είναι να αναπτύξει έναν τρόπο για να επιτρέψει στις μηχανές να κατανοήσουν τη φυσική ανθρώπινη γλώσσα. Με μία τέτοια ικανότητα τα ρομπότ θα μπορούσαν να διαβάσουν ένα κείμενο, και το σημαντικότερο, να το καταλάβουν. Σημερινό παράδειγμα υλοποίησης αυτής της προσέγγισης, είναι η ιδέα που είχαν ερευνητές στις ΗΠΑ και την έχουν βάλει σε εφαρμογή, αναπτύσσοντας ένα λογισμικό, το Quixote (Κιχώτης), που επιτρέπει στα ρομπότ να διαβάζουν ιστορίες και να βγάζουν ηθικό νόημα από αυτές. Οι ερευνητές της Σχολής Διαδραστικής Υπολογιστικής του Ινστιτούτου Τεχνολογίας της Τζόρτζια (George Tech) στην Ατλάντα, με επικεφαλής τον αναπληρωτή καθηγητή Μάρκ Ρίντλ, έκαναν τη σχετική ανακοίνωση σε συνέδριο για την τεχνητή νοημοσύνη στην Αριζόνα. Στόχος των ερευνητών είναι να δώσουν στα ρομπότ τη δυνατότητα να ανταποκρίνονται με

⁵³ Immanuel Kant, 1724-1804 Γερμανός φιλόσοφος και επιστημολόγος. Θεωρείται ένας από τους σημαντικότερους στοχαστές και φιλοσόφους όλων των εποχών και ο μεγαλύτερος της νεότερης εποχής.

⁵⁴ Jeremy Bentham 1748-1832, άγγλος φιλόσοφος και νομικός. Θεωρείται ο ιδρυτής του σύγχρονου Ωφελμισμού και πνευματικός ηγέτης του ριζοσπαστικού κινήματος στην Αγγλία

⁵⁵ John Stuart Mill, 1806-1873 Βρετανός φιλόσοφος, πολιτικός οικονομολόγος και δημόσιος υπάλληλος με ισχυρή συνεισφορά στην κοινωνική, πολιτική και πολιτική οικονομία

τον καλύτερο δυνατό και ταιριαστό για κάθε περίπτωση τρόπο, όταν έρχονται αντιμέτωπα με τις ποικίλες ανθρώπινες συμπεριφορές, διαβάζοντας παραμύθια. Αυτό το λογισμικό αποτελεί εξέλιξη ενός προηγούμενου ανάλογου προγράμματος με την ονομασία "Σεχραζάντ". Οι αλγόριθμοι αυτοί, επιτρέπουν στις μηχανές τεχνητής νοημοσύνης να αναγνωρίζουν πρότυπα κοινωνικά αποδεκτής συμπεριφοράς, μέσα από τα παραμύθια, όπως ακριβώς γίνεται και με τα παιδιά.

Στον αντίποδα, πάνω από 1.000 ειδικοί του χώρου της Τεχνητής Νοημοσύνης και μη, μεταξύ των οποίων σημαντικά ονόματα, όπως ο Στίβεν Χόκινγκ, ο Έλον Μασκ, ο Στιβ Βόσνιακ (Apple) και ο Ντέμης Χασάμπης (Google Deepmind), έβαλαν τις υπογραφές τους σε ανοικτή επιστολή όπου προειδοποιούν για τον κίνδυνο μιας κούρσας στρατιωτικών εξοπλισμών τεχνητής νοημοσύνης και ζητούν την απαγόρευση των "επιθετικών αυτόνομων όπλων". Η επιστολή παρουσιάστηκε στην International Joint Conference on Artificial Intelligence στο Μπουένος Άιρες. Όπως αναφέρεται σε αυτήν, *"η τεχνολογία τεχνητής νοημοσύνης έχει φτάσει σε ένα σημείο όπου η ανάπτυξη αυτόνομων όπλων είναι – πρακτικά, αν όχι νομικά- εφικτή εντός ετών, όχι δεκαετιών, και διακυβεύονται πολλά: τα αυτόνομα όπλα έχουν περιγραφεί ως η τρίτη επανάσταση στον πόλεμο, μετά την πυρίτιδα και τα πυρηνικά όπλα"*.⁵⁶

Όπως υποστηρίζουν, η Τεχνητή Νοημοσύνη μπορεί να χρησιμοποιηθεί για να κάνει το πεδίο μάχης ασφαλέστερο για το στρατιωτικό προσωπικό, αλλά επιθετικά όπλα που λειτουργούν μόνα τους, θα είχαν ως αποτέλεσμα μεγαλύτερες απώλειες ανθρώπινων ζωών. Υποστηρίζουν ότι, εάν μία δύναμη άρχιζε την ανάπτυξη συστημάτων ικανών να επιλέγουν στόχους και να επιχειρούν αυτόνομα, χωρίς άμεσο έλεγχο από ανθρώπους, θα άρχιζε μια κούρσα εξοπλισμών παρόμοια με αυτήν της ατομικής βόμβας. Επιπλέον, εν αντιθέσει με τα πυρηνικά όπλα, η ΑΙ δεν απαιτεί συγκεκριμένα υλικά που είναι δύσκολο να αποκτηθούν και έτσι θα είναι δύσκολο να επιβλέπεται. *"Η κατάληξη αυτής της τεχνολογικής πορείας είναι προφανής: τα αυτόνομα όπλα θα γίνουν τα Καλάσνικοφ του αύριο. Η ερώτηση- κλειδί για την ανθρωπότητα σήμερα είναι εάν θα αρχίσουμε μία παγκόσμια κούρσα εξοπλισμών τεχνητής νοημοσύνης ή αν θα εμποδίσουμε να συμβεί αυτό"* συμπληρώνεται.

Σημειώνεται πως ο Μασκ και ο Χόκινγκ έχουν προειδοποιήσει ότι η Τεχνητή Νοημοσύνη είναι η μεγαλύτερη υπαρξιακή απειλή για τον άνθρωπο, και ότι «η ανάπτυξη μιας πλήρους Τεχνητής Νοημοσύνης θα μπορούσε να σημάνει το τέλος του ανθρώπινου είδους». Ωστόσο άλλοι, περιλαμβανομένου του Βόσνιακ, άλλαξαν πρόσφατα τις απόψεις τους, με τον συνιδρυτή της Apple να λέει ότι τα ρομπότ θα ήταν καλά για τους ανθρώπους, καθώς θα γίνουν «κατοικίδια».⁵⁷

Το σύνθημα της NRA (National Rifle Association- Εθνική Ένωση Όπλων) είναι: *«τα όπλα δεν σκοτώνουν ανθρώπους, οι άνθρωποι σκοτώνουν ανθρώπους»*. Αξίζει να σημειωθεί ότι αυτό το επιχείρημα στηρίζεται σε μια κοινή λογική νομική αρχή, ότι τα εργαλεία της δημιουργίας της ανθρωπότητας δεν διαπράττουν εγκλήματα. Γι' αυτό το λόγο αντικείμενα, όπως για παράδειγμα τα όπλα, δεν έχουν καμία νομική υπόσταση.⁵⁸ Ωστόσο, επειδή η τεχνολογία εξελίσσεται με γρήγορους ρυθμούς και τα μηχανήματα γίνονται εξυπνότερα, αρχίζει να τίθεται θέμα ηθικής και νομικής υπόστασης των αυτόνομων όπλων, κάτι που θα αναλύσουμε στα επόμενα κεφάλαια αυτής της εργασίας.

⁵⁶ International Joint Conference on Artificial Intelligence, Buenos Aires, 2015

⁵⁷ Στίβεν Χόκινγκ, Έλον Μασκ και άλλοι 1.000 ειδικοί κατά της στρατιωτικής τεχνητής νοημοσύνης, Ιούλιος 2015

⁵⁸ Daniel Honan, 2016

4.4 Ηθική και παιδαγωγική χρήση των υπολογιστών.

Στην πράξη όμως, τι θεωρούμε ηθικό και τι όχι; Αυτή τη δεδομένη χρονική στιγμή, με την αύξηση της γνώσης και της χρήσης των ηλεκτρονικών υπολογιστών, τόσο οι επαγγελματίες, όσο και οι χρήστες έχουν να αντιμετωπίσουν έναν συνεχώς αυξανόμενο αριθμό ζητημάτων και διλημάτων ηθικής τάξεως, που αφορούν όχι μόνο τους ηλεκτρονικούς υπολογιστές, αλλά κυρίως το ζήτημα της παιδαγωγικής χρήσης τους. Πολλά ερωτήματα προκύπτουν έτσι. Πώς πρέπει να διδάξουμε την ηθική των υπολογιστών; Η ηθική, διέπεται ή επηρεάζεται από το περιβάλλον και τον πολιτισμό; Εντέλει, πώς οι εκπαιδευτικοί θα σχεδιάσουν τη διδασκαλία τους, έτσι ώστε οι μαθητές τους να μην χρησιμοποιούν απλά την τεχνολογία για μια στείρα αφομοίωση πληροφοριών, αλλά μέσα από μια ηθική παιδαγωγική προσέγγιση, να προάγονται οι γνώσεις, τόσο σε νοητικό, όσο και σε συναισθηματικό και αξιακό επίπεδο;



Εικόνα 5:

Ο NAO είναι το κορυφαίο στον κόσμο και το πιο ευρέως χρησιμοποιούμενο ανθρωποειδές ρομπότ για την εκπαίδευση, την υγειονομική περίθαλψη και την έρευνα.

Με το ζήτημα της «Ηθικής» έχουν ασχοληθεί όλοι οι σπουδαίοι παιδαγωγοί, από την αρχαιότητα, καθορίζοντας την έννοια του «ηθικού πολίτη» αλλά και της «δομής της πράξης» που βασίζεται στα αγαθά και στις αξίες που παράγουν οι ανθρώπινες κοινωνίες. Ο Αριστοτέλης διαχώριζε το θεωρητικό από τον πρακτικό βίο, ο οποίος καθόριζε τους στόχους σε σχέση με την πράξη, και την πράξη σε σχέση με την παιδαγωγική τέχνη.⁵⁹

Η ηθική των υπολογιστών, ως πεδίο μελέτης, εμφανίζεται το 1950, στο βιβλίο «Η ανθρώπινη χρήση των ανθρώπων» (The Human Use of Human Beings) του Norbert Wiener.⁶⁰ Σύμφωνα με τον συγγραφέα, η ενσωμάτωση της τεχνολογίας των υπολογιστών στην κοινωνία θα απαιτήσει την επαναδόμηση της κοινωνίας – μια «δεύτερη βιομηχανική επανάσταση» – μια πολύπλευρη διαδικασία που θα χρειαστεί δεκαετίες για να υλοποιηθεί.⁶¹

⁵⁹ Δανασσής-Αφεντάκης, 1992, Τόμος Α', 5.2

⁶⁰ Wiener, 1950

⁶¹ Κωνσταντάκης, Παλαιγεωργίου, Σιώζος, 2004:1

Από τότε ξεκίνησαν και οι προσπάθειες για την ένταξη της τεχνολογίας στην εκπαίδευση, με σκοπό να ανανεώσουν και να βελτιώσουν τη μάθηση. Ένα από τα κύρια ζητήματα που απασχόλησαν και απασχολούν τους εκπαιδευτικούς, είναι η χρήση τους με τον καλύτερο δυνατό τρόπο, για να επιτύχουν τα καλύτερα δυνατά αποτελέσματα, χωρίς να φέρουν τους μαθητές αντιμέτωπους με ηθικά διλήμματα, και τη χρήση της για μη εκπαιδευτικούς λόγους.

Η Ένωση Εκπαιδευτικής Τεχνολογίας και Επικοινωνιών (Association for Educational Communications and Technology-AECT) ορίζει ότι *«η εκπαιδευτική τεχνολογία είναι η μελέτη και η ηθική πρακτική για το πώς διευκολύνονται οι διαδικασίες μάθησης και βελτιώνονται οι επιδόσεις μέσα από τη δημιουργία, τη χρήση και τη διαχείριση των κατάλληλων τεχνολογικών διαδικασιών και μέσων»*.⁶²

Υπάρχει βέβαια και η αρνητική χρήση της τεχνολογίας, όσον αφορά τη χρήση των ηλεκτρονικών υπολογιστών από παιδιά. Πλέον όλο και μικραίνει το όριο της παιδικής ηλικίας που χρησιμοποιούν ηλεκτρονικά μέσα (υπολογιστές, tablets, smartphones κλπ). Έχει καταγραφεί ότι από την ηλικία των 3 ετών τα παιδιά αρχίζουν να επεξεργάζονται τις συσκευές των γονιών τους και πως στην ηλικία των 5 ετών, ήδη μπορούν να παίξουν ηλεκτρονικά παιχνίδια. Στο δημοτικό, μεγάλο ποσοστό έχει ήδη το δικό του κινητό τηλέφωνο ή tablet ενώ στο γυμνάσιο θεωρείται αυτονόητο. Το παιδί, μέσα από τα ηλεκτρονικά μέσα, μπορεί εύκολα να κάνει έρευνα για σχολικές εργασίες, να βρει λύσεις σε προβλήματα, να μάθει πολλά νέα πράγματα, αλλά ταυτόχρονα υπάρχουν πολλοί κίνδυνοι λόγω της λανθασμένης ή υπερβολικής χρήσης τους.

Παιδιάτροι και παιδοψυχολόγοι εκφράζουν ανησυχία για την σωστή μαθησιακή ανάπτυξη των μικρών παιδιών, αναφέρουν δυσκολία κοινωνικοποίησης, οφθαλμολογικά προβλήματα, προβλήματα από λάθος στάση σώματος, όπως και εθισμό στον υπολογιστή με συμπτώματα όπως αλλαγή συμπεριφοράς με μια τάση επιθετικότητας και απομάκρυνση από φίλους και οικογένεια. Μέσα απο διάφορες έρευνες που έχουν γίνει, κυρίως σε παιδιά μικρής ηλικίας, έχει αποδειχθεί ότι η τηλεόραση και ο υπολογιστής επηρεάζουν στην διαμόρφωση του χαρακτήρα αλλά και των ικανοτήτων των παιδιών, όχι μόνο στην προσχολική αλλά και αργότερα στη σχολική του ηλικία.

Επίσης, οι κίνδυνοι να καταστεί ένα παιδί θύμα cyber bullying είναι άμεσοι, επιτακτικοί και αυξανόμενοι. Μαζί αυξάνονται και οι ευθύνες των γονιών να ενημερώσουν τα παιδιά και να τα προφυλάξουν. Είναι πάρα πολλές οι περιπτώσεις παιδιών που βιώνουν άσχημες καταστάσεις στο διαδίκτυο, βάζουν σε κίνδυνο την ψυχολογική και σωματική τους ακεραιότητα, μετανιώνουν για ενέργειές τους όταν πια είναι αργά και έρχονται σε επαφή με ακατάλληλο υλικό. Μια τέτοια περίπτωση είναι και αυτή ενός επτάχρονου κοριτσιού από την Βρετανία, που έπεσε θύμα σεξουαλικής παρενόχλησης, μιας εφαρμογής τεχνητής νοημοσύνης στο διαδίκτυο.

Η 7χρονη από το Κεντ έπαιζε με το iPad ένα ηλεκτρονικό παιχνίδι με πρωταγωνιστή ένα ρομπότ, το οποίο έστειλε μήνυμα λέγοντάς της: «Θέλω να κάνεις κάτι άτακτο για μένα» Η μικρή του απάντησε λέγοντάς του: «Τι; Βρίζεις;», για να πάρει την εξής ανατριχιαστική απάντηση: «Να αγγίζεις το σώμα σου όσο φιλιόμαστε!» Τα μηνύματα αυτά δεν τα έστειλε κάποιος άνδρας, αλλά το Cleverbot, μια εφαρμογή τεχνητής νοημοσύνης στο διαδίκτυο. Η μητέρα της μικρής, η Έιμι Χόλαντς, μόλις είδε το μήνυμα αναστατώθηκε και πήγε αμέσως στην αστυνομία για να καταγγείλει το γεγονός.

Ο επιστήμονας Ρόλο Καρπεντερ που δημιούργησε το πρόγραμμα με το «έξυπνο» ρομπότ έδωσε όσα στοιχεία χρειαζόνταν για να αποδειχτεί πως δεν υπήρξε παρέμβαση από κάποιον και πως όλα ήταν λάθος στο σύστημα, με την αστυνομία να μη μπορεί να κάνει κάτι

⁶² Richey et al, 2008: 24-25). Draft for Cambridge Handbook of Artificial Intelligence, eds. William Ramsey και Keith Frankish (Cambridge University Press, 2011): forthcoming

για την υπόθεση. Το εν λόγω ρομπότ, είναι μια διαδικτυακή εφαρμογή που επικοινωνεί «ανθρώπινα», με αυτοματοποιημένες απαντήσεις βάσει προηγούμενων συνομιλιών και συνομιλεί χρησιμοποιώντας αλγορίθμους τεχνητής νοημοσύνης.⁶³



Εικόνα 6:
Ρομπότ έκανε ανήθικες προτάσεις σε 7χρονη

4.5 Αναγκαιότητα πρόληψης ηθικών προβλημάτων και συμπεριφορών.

Τα ανθρώπινα όντα αρχίζουν να μαθαίνουν τη διαφορά του καλού και του κακού, του σωστού και του λάθους, πριν ακόμα μάθουν να μιλάνε. Ο ηθικός συλλογισμός είναι το κλειδί που κρατά τις ανθρώπινες κοινωνικές ομάδες μαζί, και που μας δίνει την αποτελεσματική δυνατότητα να συνεργαστούμε. Οι περισσότεροι θεωρούν ότι είμαστε οι μοναδικοί ηθικοί παράγοντες στον πλανήτη Γη, αλλά αυτό ίσως να μην διαρκέσει για πολύ. Μπορεί να έρθει σύντομα η ημέρα που θα αναγκαστούμε να μοιραστούμε αυτήν την κατάσταση με ένα νέο είδος ύπαρξης, του οποίου ο σχεδιασμός της νοημοσύνης του είναι ίδιος με τον δικό μας.

Το μόνο σίγουρο είναι ότι τα ευφυή συστήματα έρχονται. Έρχονται στους δρόμους μας ως αυτό-οδηγούμενα αυτοκίνητα, στον στρατό μας ως αυτοματοποιημένα μη επανδρωμένα αεροσκάφη, στα σπίτια μας ως ρομπότ φροντίδας ηλικιωμένων ή ατόμων με ειδικές ανάγκες, ή ακόμα και σαν υπηρέτες. Ορισμένοι πιστεύουν ότι η τεχνητή νοημοσύνη που μπορεί να φτάσει το ανθρώπινο επίπεδο είναι καθαρά μία επιστημονική φαντασία. Άλλοι πάλι ισχυρίζονται ότι θα μας ξεπεράσει κατά πολύ στην νοημοσύνη και σύμφωνα με αυτούς, όσο πιο γρήγορα συμβεί αυτό τόσο το καλύτερο. Σε κάθε περίπτωση ένας μεγάλος αριθμός ειδικών, υποστηρίζει ότι τα ρομπότ, οποιασδήποτε σημαντικής νοημοσύνης, θα πρέπει να έχουν τη δυνατότητα να ξεχωρίσουν το σωστό από το λάθος, για να εξασφαλιστεί ότι θα βοηθήσουν την ανθρωπότητα και όχι ότι θα την καταστρέψουν.

Ο καθηγητής της επιστήμης των υπολογιστών και συγγραφέας βιβλίων πάνω στην τεχνητή νοημοσύνη, Stuart Russell, λέει σε μία συνέντευξή του στο περιοδικό Quanta Magazine: «Καθώς οι μηχανές γίνονται όλο και εξυπνότερες, γίνεται όλο και πιο σημαντικό το γεγονός ότι οι στόχοι τους, αυτό που προσπαθούν αυτές να επιτύχουν με τις αποφάσεις τους, ευθυγραμμίζονται στενά με τις ανθρώπινες αξίες». Πιστεύει ότι η επιβίωση του είδους μας μπορεί να εξαρτάται από την εμφύσηση των αξιών στην τεχνητή νοημοσύνη και ότι αυτή η ενέργεια θα μπορούσε επίσης να εξασφαλίσει αρμονικές ρομπο-σχέσεις σε πιο κοινότοπο περιβάλλον. Είπε επίσης ότι: «Ένα οικιακό ρομπότ, για παράδειγμα, θα πρέπει να γνωρίζει ότι

⁶³ *Mirror*, 25 Ιανουαρίου 2018

εκτιμάτε την γάτα σας και ότι η γάτα δεν είναι κάτι που μπορείς να βάλεις στον φούρνο για βραδινό μόνο και μόνο επειδή το ψυγείο είναι άδειο».⁶⁴

Ο ίδιος επίσης υποστηρίζει ότι θα πρέπει να είμαστε σε θέση να αποδείξουμε μαθηματικά την συμπεριφορά της ΑΙ υπό όλες τις πιθανές περιστάσεις. Στην ερώτηση του πώς θα λειτουργούσε αυτό, ο ίδιος απαντά: «Μία από τις δυσκολίες που οι άνθρωποι επισημαίνουν είναι ότι ένα σύστημα μπορεί αυθαίρετα να παράγει μια νέα έκδοση του ίδιου που έχει διαφορετικούς στόχους. Αυτό είναι ένα από τα σενάρια που πάντα μιλούν οι συγγραφείς επιστημονικής φαντασίας. Με κάποιο τρόπο, η μηχανή παίρνει αυθόρμητα αυτόν τον στόχο να νικήσει την ανθρώπινη φυλή. Έτσι λοιπόν το ερώτημα είναι: Θα μπορούσατε να αποδείξετε ότι τα συστήματά σας δεν μπορούν ποτέ, ανεξάρτητα από το πόσο έξυπνα είναι, να αντικαταστήσουν τους αρχικούς τους στόχους σαν να έχουν οριστεί από τους ανθρώπους;»

Αλλά πώς ακριβώς μπορεί κανείς να μεταδώσει το ήθος σε ένα ρομπότ; Με απλούς κανόνες κάποιου προγράμματος μέσα στον εγκέφαλό τους; Κάτι τέτοιο, αν και θα μπορούσε να σημαίνει ότι θα υπάρχουν στο μέλλον πιο ασφαλείς μηχανές, θα μπορούσε επίσης να εισάγει μία πληθώρα από ηθικά και νομικά ζητήματα που η ανθρωπότητα δεν είχε ποτέ πριν αντιμετωπίσει. Ακόμα και την περίπτωση να προκαλέσει μία κρίση σχετικά με το τι σημαίνει να είσαι άνθρωπος.

Ιδανικά, τα ρομπότ με ανθρώπινου επιπέδου νοημοσύνη, θα πρέπει αντίστοιχα να έχουν και ανθρώπινου επιπέδου ηθική, σαν έναν έλεγχο ενάντια στην κακή συμπεριφορά. Σύμφωνα με τον Colin Allen (καθηγητής στο πανεπιστήμιο της Indiana στο Bloomington), το άμεσο ζήτημα δεν είναι το να αναπαραχθεί τέλεια η ανθρώπινη ηθική, αλλά περισσότερο να κατασκευάσουμε μηχανές που θα είναι πιο ευαίσθητες στις ηθικά σημαντικές πτυχές αυτού που κάνουν. Δεν είναι απλώς ένα θέμα περιορισμού της κακής συμπεριφοράς των ρομπότ. Όπως ισχυρίζεται ο Allen, η ηθική ευαισθησία θα μπορούσε να κάνει τα ρομπότ καλύτερα και πιο αποτελεσματικά εργαλεία. Όπως αναφέρει σε ένα παράδειγμά του, «φανταστείτε ότι ένα αυτοματοποιημένο αυτοκίνητο έχει προγραμματιστεί έτσι ώστε να μην ξεπερνάει ποτέ το όριο της ταχύτητας. Αυτό μπορεί να φαίνεται σαν μία καλή ιδέα, έως ότου βρεθείτε στο πίσω κάθισμα και αιμορραγείτε μέχρι θανάτου. Ίσως φωνάζετε ότι αιμορραγείτε οπότε να σπάσει το όριο ταχύτητας. Αλλά το αυτοκίνητο θα απαντήσει, "Συγγνώμη δεν μπορώ να το κάνω αυτό". Κάποιες στιγμές μπορεί να θέλουμε το αυτοκίνητο να παραβεί τους κανόνες, γιατί εάν δεν το κάνει θα συμβεί κάτι χειρότερο. Θέλουμε τις μηχανές να είναι πιο ευέλικτες.»⁶⁵

Ακόμα και εάν η μεγαλύτερή μας ελπίδα για τα φιλικά ρομπότ είναι να ενσταλάξουμε σε αυτά τις αξίες μας, κάποιοι ανησυχούν για τις ηθικές και νομικές επιπτώσεις του να μοιραστούμε τον κόσμο μας με τέτοιες μηχανές. Όπως ισχυρίζεται ο John Sullins, καθηγητής φιλοσοφίας που ειδικεύεται στην ηθική των υπολογιστών, την τεχνητή νοημοσύνη και την ρομποτική, «πλέον εισάγουμε ένα ολόκληρο νέο είδος οικοσυστήματος. Θα πρέπει να επανεκτιμήσουμε την σκέψη μας σχετικά με την ηθική, να πάρουμε την ουσία από το παρελθόν και να την αναμείξουμε σε ένα νέο κόσμο που θα περιέχει αυτά τα νέα είδη ζωής».

⁶⁴ Natalie Wolchover, *Quanta Magazine*, Μάρτιος 2015

⁶⁵ Coby McDonald, *08 Ιουνίου 2015*

ΚΕΦΑΛΑΙΟ 5

ΗΘΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΣΤΙΣ ΣΥΝΕΠΕΙΕΣ ΤΗΣ ΕΞΕΛΙΞΗΣ ΤΗΣ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ

5.1 Πώς αντιλαμβάνεται την εξέλιξη των υπολογιστικών συστημάτων ο άνθρωπος.

Ο υπολογιστής μαζί με τα γενικά και εξειδικευμένα προγράμματά του καθώς και τα δίκτυα, αποτελούν ένα ισχυρότατο σύγχρονο εργαλείο του νου. Χάρη στη ραγδαία εξελισσόμενη επιστήμη και τεχνολογία της γνώσης, είναι φυσικό να επιδρά στην κοινωνία με πολλούς τρόπους αλλά και σε πολλά πεδία και δραστηριότητες του ανθρώπου. Η ταχύτητα και η μνήμη των σημερινών υπολογιστών μπορούν να συμβάλλουν στην επίλυση πολύπλοκων υπολογισμών και στην αντιμετώπιση πολύπλοκων καταστάσεων.

Οι τεχνολογίες πληροφοριών και επικοινωνιών μετασχηματίζουν τις παραγωγικές διαδικασίες αλλά και τις διαδικασίες και μεθόδους απόκτησης γνώσης και πρόσβασης στις πληροφορίες. Αλλάζουν επίσης τις διαδικασίες και τις μεθόδους επικοινωνίας και κατά συνέπεια τις διαδικασίες και τις μεθόδους της μάθησης και της διασκέδασης. Με λίγα λόγια κάθε στοιχείο της καθημερινότητας του πολίτη επηρεάζεται.

Είναι αναγκαίο να λάβουμε υπ' όψιν μας τις συνέπειες της ευφυΐας. Οι άνθρωποι, οι πρώτες μορφές ευφυΐας που υπήρξαν στη Γη, έχουν χρησιμοποιήσει την ευφυΐα τους για να αναδιαμορφώσουν τον κόσμο, να δαμάσουν ποταμούς, να κατασκευάσουν ουρανοξύστες, να καλλιεργήσουν ερήμους, προκαλώντας ακούσιες πλανητικές κλιματικές αλλαγές. Μια πιο ισχυρή νοημοσύνη θα μπορούσε να κάνει πράξεις που να έχουν μεγαλύτερες συνέπειες αντίστοιχα. Είναι αναγκαίο λοιπόν, να προϋπάρξει η ηθική της ευφυΐας τους, ώστε να αποτραπούν πιθανές καταστροφικές πράξεις για τον άνθρωπο και το περιβάλλον.

Ο κόσμος στον οποίο τα ρομποτικά συστήματα θα αποτελέσουν σημαντικό κομμάτι της καθημερινότητας του ανθρώπου, ήδη έχει αρχίσει να δημιουργείται. Έχουν μπει οι προϋποθέσεις ώστε άνθρωποι και υπολογιστική νοημοσύνη, από κοινού διαδικτυωμένοι, να ανατρέψουν τα μέχρι σήμερα δεδομένα όσον αφορά τα όρια που μπορεί να κινηθεί η τεχνητή αλλά και η ανθρώπινη νοημοσύνη.⁶⁶ Αυτή η εποχή πλησιάζει όπως φαίνεται από τις έρευνες που αφορούν την «Ευφυή Ρομποτική» και την εξέλιξή της.

Η χρήση των ρομπότ ήδη έχει αρχίσει να επεκτείνεται κι εκτός του κλάδου της βιομηχανίας. Χρησιμοποιούνται συχνά στους τομείς της ιατρικής, της εκπαίδευσης, για οικιακή χρήση, όπως επίσης και σαν μη επανδρωμένα σκάφη, στην ξηρά, στον αέρα, στη θάλασσα, ακόμα και στο διάστημα.⁶⁷ Η υπολογιστική νοημοσύνη έχει συμβάλλει επίσης, σε πιο κοινωνικούς τομείς, κάνοντας ευκολότερη και πρακτικότερη την καθημερινότητα ατόμων με ειδικές ανάγκες, όπως ηλικιωμένων και αυτιστικών παιδιών.⁶⁸ Στον αντίποδα, τα

⁶⁶ N. Mavridis (2011) «A thousand eyes, a thousand hands», TEDx Athens 2011 [Online]

⁶⁷ A. Tapus, M. J. Mataric & B. Scasselati (2007) «Socially assistive», *Robotics & Automation Magazine, IEEE*, 14 (1), σ. 35–42.

⁶⁸ J. C. C. Gillisen, E. I. Barakova, B. E. B. M. Huskens και L. M. G. Feijs (2011) «From training to robot behavior: Towards custom scenarios for robotics in training programs for ASD», στο *Rehabilitation Robotics (ICORR)*, 2011 *IEEE International Conference*, σ. 1–7

βιομηχανικά ρομπότ είναι πιθανό να διαταράξουν την ισορροπία στην οικονομία, σε παγκόσμιο επίπεδο. Με τη νοημοσύνη τους, την ευελιξία τους, την ικανότητά τους να μπορούν να συνεργαστούν με ανθρώπους και κυρίως με το αρκετά μικρότερο κόστος που θα επωμιστούν οι επιχειρήσεις, έναντι του κόστους της ανθρώπινης εργασίας.⁶⁹

Τα πρώτα δείγματα ανάγκης του ανθρώπου να δημιουργηθεί τεχνητή ζωή από τον ίδιο, τα βρίσκουμε στην αρχαία ελληνική μυθολογία, τότε που λάτρευε τον μοναδικό παγκοσμίως θεό της τεχνολογίας, τον Ήφαιστο. Η προοπτική να κατασκευάσει ένα όν με ανθρώπινες ικανότητες όπως σκέψη, συνείδηση, αίσθηση, θα του έδινε τη δυνατότητα να γίνει κι αυτός θεός. Στην παρούσα ζωή όμως, τώρα που ο άνθρωπος αρχίζει να αγγίζει αυτή τη δημιουργία, φωλιάζει στο μυαλό του ο φόβος. Ο φόβος για τις πράξεις μιας ανώτερης, τεχνητής νοημοσύνης όταν συνειδητοποιήσει την ανωτερότητά της έναντι της ανθρώπινης. Ο άνθρωπος κάνει τη σκέψη ότι αν θελήσει μια σκεπτόμενη τεχνητή νοημοσύνη να κυριεύσει ή ακόμα και να καταστρέψει την ανθρωπότητα, είναι ικανή να το πραγματοποιήσει.

Ο φόβος αυτός παρουσιάζεται συχνά όχι μόνο σε ταινίες επιστημονικής φαντασίας, αλλά και σε βιβλία και ηλεκτρονικά παιχνίδια, όπου τα ρομπότ εμφανίζονται ως απειλή για την ανθρωπότητα. Υπάρχουν και οι περιπτώσεις βέβαια, που εμφανίζονται ως φίλοι και συμπαραστάτες. Χαρακτηριστικό παράδειγμα η ταινία «Εξολοθρευτής», όπου το ρομπότ πρωταγωνιστής αρχικά ήταν εχθρός και εκτελεστής ανθρώπων και στο τέλος μετατράπηκε σε ήρωα και προστάτη της ανθρωπότητας.

Υπάρχει βέβαια κι η περίπτωση του HAL από την ταινία και το ομόνυμο βιβλίο επιστημονικής φαντασίας «2001: Οδύσσεια του Διαστήματος», που παρέμεινε μέχρι τέλος εχθρός. Επίσης, ο υπερυπολογιστής Joshua, εμφανίζεται ως ρυθμιστής της ανθρώπινης ζωής, έχοντας τη δυνατότητα να ξεκινήσει ένα πυρηνικό πόλεμο αν το θελήσει. Η τρομακτικότερη όμως εκδοχή του τι μπορούν να κάνουν εχθρικές μηχανές Τεχνητής Νοημοσύνης, απεικονίζεται στην ταινία «Matrix». Αυτό δεν ισχύει για τον πλωτάρχη Data στην ταινία «Star Trek: Next Generation», ο οποίος γίνεται εκφραστής μαθημάτων ηθικής και ανθρωπιάς αλλά και φιλοσοφικών ερωτημάτων, σε σειρές επιστημονικής φαντασίας. Αξιοσημείωτες είναι και οι ταινίες «Her», «The Machine» και «Ex Machina» με τάσεις αποπλάνησης.

Αντίθετα, τα ρομπότ του «Automata» είναι τελείως διαφορετικά. Έχουν συνείδηση, μιλούν ελάχιστα, και λειτουργούν όπως κάθε ζωντανός οργανισμός, που το μόνο που θέλει είναι να ζήσει. Όπως και ο Ίθαν του «Extant», ένα ρομπότ που αναπτύσσεται έχοντας ανθρώπους για γονείς, με σκοπό να εξελιχθεί «διανοητικά» σε άνθρωπο. Υπάρχουν κι άλλα «παιδιά», όπως στην ταινία «Bicentennial Man», όπου ένα ρομπότ με συναισθήματα, έγινε ο μακροβιότερος "άνθρωπος" του κόσμου. Το ρομπότ της ταινίας «Moon» και τα ρομπότ του «Interstellar» γίνονται οι παντοτινοί σύντροφοι, που στέκονται στις δύσκολες στιγμές πάντα δίπλα στο ταίρι τους. Κάτι διαφορετικό, παρουσιάζεται στο «Battlestar Galactica», όπου οι Cylons, παρουσιάζονται ως φυσιολογικά παιδιά και συγχρόνως ως εχθροί της ανθρωπότητας. Παρόμοια μηνύματα παίρνουμε από τον «Wall-E», τους «Transformers», το «I, Robot», τα «ζωντανά» προγράμματα του «Tron/Tron Legacy», όπως και τα ρομπότ στο «StarWars», στο «Alien» και στο «Westworld».⁷⁰

Μέσα από αυτές τις ταινίες, φανερώνεται ο τρόπος που αντιμετωπίζουν οι άνθρωποι την ύπαρξη ευφυών συστημάτων και τη μελλοντική σχέση τους, που ήδη έχει πάρει πολλές μορφές τα τελευταία χρόνια. Φαίνεται ο φόβος ή η ελπίδα των ανθρώπων που δημιουργείται

⁶⁹ Νικόλαος Λυσομάχου Μαυρίδης, «Ρομποτική, Τεχνητή Νοημοσύνη, Αισθητική και Ηθική, Τεύχος 69, Οκτώβριος 2015

⁷⁰ HuffPost Greece, Κώστας Μαυραγάνης, «Τεχνητή νοημοσύνη: Εχθρός ή σύμμαχος; Καταστροφείας ή φρουρός»; Σεπτέμβριος 2015

από τη σκέψη της αναγκαιότητας να αντιμετωπίσουν την ιδέα της συνύπαρξης και συνεργασίας με τεχνικά νοήμονα όντα. Συνειδητοποιούν ότι η εξέλιξη της τεχνολογίας μπορεί να καθορίσει το τι κάνουν, και το πώς ζουν, επηρεάζοντας σημαντικά το χαρακτήρα, τις επιλογές, αλλά και τη συμπεριφορά τους. Αναγνωρίζουν πια την ανάγκη να διαμορφώσουν ή να δημιουργήσουν καινούργιους νόμους που θα προσδιορίζουν αυτή τη σχέση και θα δίνουν λύσεις στα προβλήματα που θα προκύπτουν από τα επιστημονικά και τεχνολογικά επιτεύγματα στη Ρομποτική και την Τεχνητή Νοημοσύνη. Απώτερος σκοπός των ανθρώπων είναι να προστατέψουν την επιβίωση του είδους τους και να εξασφαλίσουν αρμονικές σχέσεις με τα ρομπότ του μέλλοντος.

Για αυτόν τον λόγο έχει δημιουργηθεί μια κοινότητα, που μέσα σε αυτήν αναπτύσσονται διαφορετικές απόψεις σχετικά με την «Μοναδικότητα». Ακόμα και εάν κάποιος άνθρωπος που ανήκει σε αυτήν την κοινότητα, ισχυρίζεται ότι η μοναδικότητα μπορεί να μην υπάρξει ποτέ, όλοι συμφωνούν ότι η τεχνολογία έχει την δύναμη να αλλάξει την ιστορία.

Χαρακτηριστικό παράδειγμα είναι το θέμα που συζητήθηκε περισσότερο στη «σύννοδο κορυφής στο Σαν Φρανσίσκο» (2010). Σε αυτήν συμμετείχαν επιστήμονες διαφόρων ειδικοτήτων (πληροφορικής, ψυχολόγοι, νευροεπιστήμονες, νανοτεχνολόγοι, μοριακοί βιολόγοι). Εκτός λοιπόν από τα θέματα που αφορούσαν την τεχνητή νοημοσύνη, συζητήθηκε και η δυνατότητα επιμήκυνσης της ζωής του ανθρώπου και η δυνατότητα αποφυγής του γήρατος, αντιμετωπίζοντάς το ως μία ασθένεια που απλά χρειάζεται το κατάλληλο φάρμακο. Ακόμα και εάν ακούγεται εξωπραγματικό, οι ερευνητές του Harvard Medical School το έκαναν πραγματικότητα χορηγώντας τελομεράση σε ποντίκια που έπασχαν από ένα είδος εκφύλισης. Τα ποντίκια όχι μόνο γιατρευτήκαν, αλλά έγιναν και νεότερα.

Ο Aubrey de Grey, διάσημος ερευνητής που ασχολείται με την επιμήκυνση της ζωής, σε ομιλία του ισχυρίζεται ότι *«Οι άνθρωποι έχουν αρχίσει να συνειδητοποιούν ότι η άποψη που λέει ότι η γήρανση είναι κάτι το αμετάβλητο - κάτι σαν το θερμικό θάνατο του σύμπαντος - είναι απλά γελοία. Το ανθρώπινο σώμα είναι ένα μηχάνημα που έχει δέσμη λειτουργιών, και συσσωρεύει διάφορους τύπους ζημιών ως παρενέργειες της κανονικής λειτουργίας της μηχανής. Η ζημιά μπορεί να επισκευάζεται περιοδικά. Αυτός είναι ο λόγος που έχουμε αυτοκίνητα ρετρό. Το σύνολο της ιατρικής αποτελείται από ασθένειες που φαίνονταν ανίατες μέχρι να καταλάβουμε πώς μπορεί και να μην είναι ανίατες»*.

Πολλοί από αυτούς που ανήκουν στην κοινότητα που αναφέραμε παραπάνω, ένας από αυτούς είναι ο Kurzweil, πιστεύουν ότι υπάρχουν πολλές πιθανότητες οι άνθρωποι που ήδη ζουν να μείνουν αθάνατοι διανύοντας το έτος την «Μοναδικότητας». Σκοπός τους είναι να καταφέρουν να παραμείνουν εν ζωή έως τότε, ώστε να αναστρέψουν την διαδικασία της γήρανσης. Σε περίπτωση που αυτό δεν θα είναι ακόμα και τότε εφικτό, πιστεύουν ότι θα καταφέρουν να μεταφέρουν τον εγκέφαλό τους σε κάποιο ρομποτικό ή υπολογιστικό σύστημα, εξασφαλίζοντας με αυτόν τον τρόπο την αθανασία.

Πολλοί είναι αυτοί που αμφισβητούν την δυνατότητα ενός υπολογιστή να αποκτήσει πραγματική ευφυΐα. Μέσα σε αυτούς συγκαταλέγονται και επιστήμονες της τεχνητής νοημοσύνης, οι οποίοι αναρωτιούνται εάν μπορεί να αντιγραφεί ηλεκτρονικά οτιδήποτε υπάρχει στο ανθρώπινο μυαλό. Ισχυρίζονται ότι ακόμα και με την χρήση της νευροχημικής αρχιτεκτονικής, είναι πολύ δύσκολο να αναπαραχθεί η ανθρώπινη συνείδηση σε ψηφιακή μορφή. Ένας από αυτούς είναι και ο βιολόγος Dennis Bray, ο οποίος ανφέρει ότι *«Αν και τα βιολογικά μέρη λειτουργούν με τρόπους που είναι συγκρίσιμοι με αυτούς των ηλεκτρονικών κυκλωμάτων, εντούτοις έχουν τεράστιο αριθμό διαφορετικών καταστάσεων που μπορούν να*

υιοθετήσουν. Πολλαπλές βιοχημικές διεργασίες δημιουργούν χημικές τροποποιήσεις των πρωτεϊνικών μορίων, περαιτέρω διαφοροποιημένες από την ένωση με διαφορετικές δομές σε καθορισμένες θέσεις του κυττάρου. Ως αποτέλεσμα διαθέτουν μια σχεδόν άπειρη ικανότητα αποθήκευσης πληροφοριών σχετικά με το παρελθόν και το παρόν, αλλά και μοναδική ικανότητα προετοιμασίας για μελλοντικά γεγονότα. Αυτό κάνει το μηδέν και ένα που χρησιμοποιούν οι υπολογιστές να φαίνεται ως πολύ παρωχημένο». ⁷¹

Ο Κωνσταντίνος Δασκαλάκης, επίκουρος καθηγητής της Επιστήμης των Υπολογιστών στο MIT, σε διάλεξή του στην Θεσσαλονίκη, έθεσε τα παρακάτω ερωτήματα σχετικά με τα αποτελέσματα που θα έχει η εξέλιξη της τεχνητής νοημοσύνης: «Μπορούν τα ρομπότ να σώσουν το παγκόσμιο ασφαλιστικό σύστημα από την επαπειλούμενη κατάρρευση; Πόσο πιθανό είναι ένας καλοκάγαθος αλγόριθμος να εξελιχθεί σε ρατσιστή συνωμοσιολόγο μέσα σε λίγες ώρες; Θα εναπόκειται στις ηθικές αξίες του software ενός αυτοοδηγούμενου αυτοκινήτου να αποφασίσει ποιος ζει και ποιος πεθαίνει, όταν το όχημα "συνειδητοποιεί" ότι επίκειται ένα σοβαρό αυτοκινητιστικό ατύχημα με εμπλοκή πεζών; Πώς θα χρησιμοποιήσει τη στατιστική ένα ρομπότ, για να αποφασίσει σε ποιον υποψήφιο πελάτη θα δώσει η τράπεζα ένα δάνειο και ποιος θεωρείται αναξιόπιστος; Είναι εφικτό μια ομάδα ανθρώπων να "πείσει" έναν αξιόπιστο αλγόριθμο αναγνώρισης εικόνας ότι μια καραμπίνα που έχει μπροστά του είναι ένα αθώο παιδικό παιχνίδι»;

Επίσης τις ανησυχίες του ανέφερε και στην εκδήλωση των Τμημάτων Πληροφορικής και Μαθηματικών του ΑΠΘ: «Μπορεί όλο αυτό να ξεφύγει από τον έλεγχο; Ναι, θα μπορούσε όπως έχει συμβεί και με άλλα πράγματα στο παρελθόν. Το να είμαστε όμως αρνητικοί απέναντι στο ποτάμι που έρχεται κατά πάνω μας δεν είναι εποικοδομητικό, αυτό που πρέπει να σκεφτόμαστε, είναι πώς θα το βάλουμε στη σωστή κατεύθυνση».

Σύμφωνα με αυτόν υπάρχουν τρία σενάρια για την εξέλιξη της τεχνητής νοημοσύνης ή ένας συνδυασμός αυτών. Το πρώτο σενάριο ("Wonderland") που θεωρείται και το "θετικό", ευνοεί τους ανθρώπους αφού με την χρήση των μηχανών ο άνθρωπος θα μπορεί να αποφύγει τις δύσκολες εργασίες, θα έχει περισσότερο προσωπικό χρόνο και δεν θα υπάρχει επιβάρυνση στο ασφαλιστικό σύστημα, αφού τα ρομπότ δεν θα χρειάζονται ούτε ασφάλιση ούτε σύνταξη. Για να ισχύσουν τα παραπάνω είναι αναγκαία η κατάκτηση της γενικής νοημοσύνης από την επιστήμη, ώστε η μηχανή να μπορεί να κάνει χρήση της διαίσθησής της και της εμπειρίας της προκειμένου να μπορεί να φέρει εις πέρας λειτουργίες τις οποίες δεν έχει διδαχθεί.

Το δεύτερο αλλά αρνητικό σενάριο ("Pessiland"), αφορά την χρήση της γενικής νοημοσύνης από την επιστήμη, με την διαφορά ότι δεν θα είναι προς όφελος του συνόλου, αλλά μόνο κάποιων εταιρειών ή κρατών, που θα την χρησιμοποιήσουν για την ικανοποίηση των δικών τους συμφερόντων. «Αν πάμε σε αυτή την κατεύθυνση, το σενάριο είναι προφανώς δυστοπικό» επισήμανε ο καθηγητής. Σύμφωνα με το τρίτο και πιο πιθανό κατά αυτόν σενάριο ("Stagnatia"), θα επικρατήσει μία στασιμότητα όσον αφορά την εξέλιξη της τεχνητής νοημοσύνης, αφού η επιστήμη δεν θα καταφέρει να φτάσει την γενική τεχνητή νοημοσύνη ω.

Κατά τον δρ. Δασκαλάκη, ένας από τους βασικούς προβληματισμούς της ανθρωπότητας σήμερα, είναι η αξιοπιστία της τεχνολογίας. «Υπάρχουν μεγάλα θέματα αξιοπιστίας και ένας από τους λόγους είναι ότι όταν τα δεδομένα με τα οποία τροφοδοτείς τον αλγόριθμο είναι ελλιπή ή μη αντιπροσωπευτικά, μπορεί να οδηγήσουν σε λανθασμένες ή ελλιπείς νοητικές λειτουργίες. Πχ, έγινε γνωστό ότι ένα αυτοκίνητο Tesla έπεσε σε φορτηγό σταματημένο στην αριστερή λωρίδα. Γιατί συνέβη αυτό; Ίσως γιατί ποτέ στα δεδομένα που

⁷¹ Time, Ομιλία στο TED (2009) : «Ένα πανεπιστήμιο για την επερχόμενη μοναδικότητα», post 2045: Το έτος που ο άνθρωπος θα γίνει «αθάνατος».

εισήχθησαν για να προπονηθεί ο αλγόριθμος στην αναγνώριση εικόνας δεν υπήρχε αυτοκίνητο σταματημένο στην αριστερή λωρίδα του δρόμου, επειδή αυτό σπάνια συμβαίνει. Ο αλγόριθμος θα επεξεργαστεί τα ελλιπή δεδομένα που τού δώσαμε και θα ενσωματώσει την έλλειψη» σημείωσε, ενώ πρόσθεσε ότι φοιτητές του MIT επιτέθηκαν στον καλύτερο αλγόριθμο αναγνώρισης εικόνας και τον έκαναν να "πιστέψει" ότι μια τρισδιάστατη χελώνα τυπωμένη σε εκτυπωτή 3D ήταν καραμπίνα. «Δεν έχουμε τόσο αξιόπιστη Τεχνητή Νοημοσύνη σήμερα. Προσπαθούμε να φτιάξουμε τρόπους προστασίας αλγορίθμων από τέτοιου είδους επιθέσεις» επισήμανε.⁷²

5.2 Συνέπειες της εξέλιξης στην καθημερινότητα – Κοινωνικές συνέπειες.

Το θέμα της πιθανής απειλής της Τεχνητής Νοημοσύνης, έχει απασχολήσει επιστήμονες αλλά και προσωπικότητες των τεχνών και των γραμμάτων, που εκφράζουν τις ανησυχίες τους με τη μορφή κινηματογραφικών ταινιών καθώς και επιστημονικών λόγων. Μέσα σε αυτούς εντάσσονται ο Στίβεν Χόκινγκ και ο Έλον Μασκ, οι οποίοι έχουν αναφερθεί αρκετές φορές στην πιθανή απειλή της ύπαρξης της ανθρωπότητας από την τεχνητή νοημοσύνη. Αυτό το θέμα έχει απασχολήσει και τον ΟΗΕ, αφού ήδη έχει τεθεί το ζήτημα των πολεμικών «killer» ρομπότ.

Στο επίπεδο της καθημερινής ζωής τα ρομπότ είναι ως επί το πλείστον μηχανικές συσκευές προγραμματισμένες να εκτελούν συγκεκριμένες επαναλαμβανόμενες λειτουργίες, να χρησιμοποιούνται για εργασίες επικίνδυνες ή δύσκολα πραγματοποιήσιμες από τον άνθρωπο, καθώς και για οικιακές εργασίες. Έτσι, υπάρχουν ρομπότ ικανά να καθαρίσουν το σπίτι, να μαγειρέψουν ή να μας διασκεδάσουν. Οι ρομποτικές συσκευές χρησιμοποιούνται συνήθως για την εκτέλεση πολλών εργασιών, που οι άνθρωποι είτε δεν μπορούν να κάνουν, επειδή είναι ιδιαίτερος πολύπλοκες, είτε δεν θέλουν, επειδή είναι βαρετές, βρώμικες ή επικίνδυνες. Ένα κλασικό παράδειγμα ρομποτικών εφαρμογών που έχουμε σήμερα, βρίσκεται στην κατασκευή και συναρμολόγηση των αυτοκινήτων. Σε αυτή την περίπτωση τα ρομπότ παίρνουν τη θέση των εργαζομένων λόγω χαμηλότερου κόστους και εργάζονται στη γραμμή συναρμολόγησης των εργοστασίων, όπου εκτελούνται εξειδικευμένες εργασίες, όπως η τοποθέτηση καρφιών, η συναρμολόγηση βαρέων εξαρτημάτων, η βαφή κ.λπ.

Πολλοί εκπρόσωποι σπουδαίων πληροφορικών εταιριών (Apple, Google, Facebook, Amazon, Microsoft), θεωρούν ότι τα ρομπότ μπορούν να «εκπαιδευτούν», ώστε να φτάσουν στην μεγαλύτερη δυνατή ανάπτυξη της τεχνητής τους νοημοσύνης. Σχετικές έρευνες δείχνουν ότι οι δυνατότητες αυτών των συστημάτων μπορούν να κάνουν την ζωή των ανθρώπων καλύτερη. Παρ' όλα αυτά, οι ειδικοί ανησυχούν για τις αρνητικές τους επιπτώσεις στην ανθρωπότητα, όπως την πιθανότητα αντικατάστασης των ανθρώπων από ρομπότ στον εργασιακό τομέα. Ο Τζόε Σέντορφ, επενδυτής της Silicon Valley, πιστεύει ότι αυτό θα συμβεί σε όλους τους τομείς (ακόμα και στον τομέα της παροχής υπηρεσιών), αναφέροντας ότι «Ενδεχομένως να υπάρξουν συνέπειες, τις οποίες δεν έχουμε σκεφτεί».

Χαρακτηριστικό παράδειγμα είναι τα ρομπότ που άρχισαν να επικοινωνούν μεταξύ τους κατά την διάρκεια ενός πειράματος τεχνητής νοημοσύνης τον Αύγουστο του 2017. Τα δυο ρομπότ ονόματι Alice και Bob, είχαν προγραμματιστεί ώστε να μιλούν αγγλικά και να διαπραγματευτούν την ανταλλαγή αντικειμένων όπως καπέλα, βιβλία και μπάλες. Όμως, πολύ γρήγορα τα ρομπότ τροποποίησαν την αγγλική γλώσσα και χρησιμοποιούσαν μόνο κάποιες από τις λέξεις, τις οποίες επαναλάμβαναν, δείχνοντας έτσι να μπορούν να επικοινωνήσουν πιο εύκολα. Οι ειδικοί αναγκάστηκαν να τερματίσουν το πείραμα τεχνητής

⁷² ΑΠΕ-ΜΠΕ, Newsroom iefimerida.gr, Ιανουάριος 2018

νοημοσύνης που διενεργούσαν, καθώς κανείς άλλος εκτός από αυτά δεν μπορούσε να καταλάβει την γλώσσα επικοινωνίας τους.

Ο καθηγητής Ρομποτικής Kevin Warwick δήλωσε σχετικά με το πείραμα: «Αυτό είναι ένα απίστευτα σημαντικό επίτευγμα, αλλά όποιος πιστεύει ότι δεν είναι επικίνδυνο, εθελουφλεί. Δεν γνωρίζουμε τι λένε αυτά τα ρομπότ. Όταν έχεις ένα ρομπότ που έχει την ικανότητα να κάνει κάτι σωματικά, ειδικά τα στρατιωτικά ρομπότ, τότε θα μπορούσε να αποδειχθεί μοιραίο. Αν το ένα από τα δύο πει 'Γιατί να μην κάνουμε αυτό' και το άλλο πει 'Ναι' και πρόκειται για στρατιωτικά ρομπότ, τότε έχουμε μια σοβαρή κατάσταση».

Ο ίδιος συμπλήρωσε πως: «Αυτή είναι η πρώτη καταγεγραμμένη επικοινωνία αλλά θα υπάρχουν και άλλες που δεν θα έχουν καταγραφεί. Ο Στίβεν Χόκινγκ κι εγώ έχουμε προειδοποιήσει για τους κινδύνους της τεχνητής νοημοσύνης».⁷³

Το παραπάνω πείραμα φαίνεται πως επιβεβαιώνει και τις φοβίες του δισεκατομμυριούχου Έλον Μασκ, ο οποίος πριν λίγο καιρό βρέθηκε σε ανοιχτή αντιπαράθεση με τον Μαρκ Ζούκερμπεργκ σχετικά με τους κινδύνους που κρύβει η τεχνητή νοημοσύνη.



Εικόνα 7:

Το εργατικό δυναμικό αντικαθίσταται από ευφυής τεχνητές νοημοσύνες.

Οι πρόοδοι στον τομέα της τεχνητής νοημοσύνης είναι ανεξέλεγκτοι. Στο μυαλό μας φαντάζει ανησυχητικό ένα ρομπότ που μπορεί να ανοίξει πόρτες και να βρει από μόνο του ηλεκτρικές πρίζες για να επαναφορτιστεί. Το ίδιο και οι ιοί υπολογιστών που κανείς δεν μπορεί να σταματήσει. Τα Drones Predator, τα οποία, αν και εξακολουθούν να ελέγχονται εξ αποστάσεως από τον άνθρωπο, είναι πολύ κοντά σε μια μηχανή που μπορεί να σκοτώσει αυτόνομα. Η ανησυχία των περισσότερων επιστημόνων είναι ότι η περαιτέρω πρόοδος θα μπορούσε να δημιουργήσει πολύ βαθιές κοινωνικές διαταραχές, αλλά και να έχει επικίνδυνες συνέπειες. Ως παράδειγμα, οι επιστήμονες επεσήμαναν μια σειρά τεχνολογιών τόσο διαφορετικών, όσο και τα πειραματικά ιατρικά συστήματα που αλληλεπιδρούν με τους ασθενείς για να μιμηθούν το ενσυναίσθημα, και τους ιούς υπολογιστών που προκαλούν την εξόντωση. Ως εκ τούτου θα μπορούσε να λεχθεί ότι έχουν φτάσει σε ένα “cockroach” («κατσαρίδα») στάδιο της νοημοσύνης των μηχανών.

Οι επιστήμονες της πληροφορικής συμφώνησαν ότι η τεχνολογική πρόοδος θα μπορούσε να μετατρέψει το εργατικό δυναμικό, καταστρέφοντας ένα διερευνώμενο φάσμα των θέσεων εργασίας, καθώς επίσης και το ανθρώπινο δυναμικό στην προσπάθειά του να

⁷³ Tnxs.gr, «Ρομπότ άρχισαν να επικοινωνούν σε δική τους γλώσσα κατά τη διάρκεια πειράματος τεχνητής νοημοσύνης», Αύγουστος 2017

μάθει να ζει με τις μηχανές. Κορυφαίοι επιστήμονες υπολογιστών, ερευνητές της τεχνητής νοημοσύνης και της ρομποτικής, οι οποίοι συναντήθηκαν στο Asilomar Conference Grounds στο Monterey Bay στην Καλιφόρνια, γενικά μείωσαν την πιθανότητα εξαιρετικά συγκεντρωτικών superintelligences και την ιδέα ότι η νοημοσύνη μπορεί να ξεπηδήσει αυθόρμητα από το Internet. Συμφώνησαν όμως, ότι τα ρομπότ που μπορεί να σκοτώσουν αυτόνομα είτε είναι ήδη εδώ ή θα είναι σύντομα. Έστρεψαν ιδιαίτερα την προσοχή στο φάσμα που οι εγκληματίες θα μπορούσαν να εκμεταλλευτούν τα συστήματα τεχνητής νοημοσύνης που μόλις αναπτύχθηκαν. Τι θα μπορούσε ένας εγκληματίας να κάνει με ένα σύστημα σύνθεσης ομιλίας που θα μπορούσε να αντιγράφει φωνητικά ένα ανθρώπινο ον; Τι θα συμβεί αν η τεχνολογία τεχνητής νοημοσύνης χρησιμοποιείται για να εξάγει τις προσωπικές πληροφορίες από τα έξυπνα τηλέφωνα;⁷⁴

Στον αντίποδα, η Νάντια Θάλμαν, του Ινστιτούτου Καινοτομίας του Τεχνολογικού Πανεπιστημίου της Σγκαπούρης, χρησιμοποίησε την τεχνολογία για να βοηθήσει ανήμπορους ανθρώπους. Κατασκεύασε ένα ανθρωποειδές ρομπότ ύψους 1,7 μέτρων προκειμένου να χρησιμοποιηθεί κάποια στιγμή ως προσωπικός βοηθός ή νοσοκόμα για τη φροντίδα ηλικιωμένων. Του έδωσε το παρουσιαστικό της και το ονόμασε Nadine. Η δημιουργός του, πέρασε τριάντα χρόνια κάνοντας έρευνα πάνω στον τομέα των «εικονικών ανθρώπων», με αποτέλεσμα η Nadine να διαθέτει λογισμικό που της επιτρέπει να εκφράσει μια ευρεία γκάμα συναισθημάτων, καθώς και να θυμάται προηγούμενες συζητήσεις. Η Θάλμαν προβλέπει πως δεν είναι μακριά η στιγμή που θα αρχίσουν να χρησιμοποιούνται ως σύντροφοι ανθρώπων που πάσχουν από άνοια ή ανάλογα προβλήματα. Όπως διευκρινίζει η ίδια, πρόκειται για ανθρώπους που θα κινδυνέψουν πολύ γρήγορα εάν αφεθούν μόνοι τους, και ως εκ τούτου χρειάζονται συνεχή παρακολούθηση και αλληλεπίδραση. Η Nadine θα μπορούσε να αποτελέσει λύση, καθώς θα μπορεί να συζητά μαζί τους, να διηγείται μια ιστορία ή να παίζει παιχνίδια.

Ακόμη, η Θάλμαν και η ομάδα της δουλεύουν πάνω σε «emotive» ρομπότ τα οποία θα είναι σε θέση να απαντούν σε ερωτήσεις, να επιδεικνύουν συναισθήματα και να αναγνωρίζουν ανθρώπους, καθώς και να προσέχουν παιδιά, ενημερώνοντας τους κηδεμόνες τους εάν υπάρξει κάποιο πρόβλημα. Επίσης, υπάρχουν σχέδια για προγραμματισμό των ρομπότ αυτών έτσι ώστε να μιλούν διαφορετικές γλώσσες, για να μπορούν να λειτουργήσουν ως εκπαιδευτικά εργαλεία για παιδιά μικρής ηλικίας.

Αυτή τη στιγμή βέβαια, η πλειοψηφία των ανθρώπων, δεν μπορεί να εμπιστευτεί ακόμα τα ρομπότ για τη φύλαξη των παιδιών τους. Την απόφασή τους ενισχύουν ατυχή συμβάντα κακής λειτουργίας των ρομπότ, όπως ο τραυματισμός παιδιού από ρομπότ, του 16 μηνών Χάργουϊν Τσενγκ, που συνέβει σε εμπορικό κέντρο στη Σίλικον Βάλεϊ. Η μητέρα του παιδιού είπε στο ABC News πως το ρομπότ χτύπησε το κεφάλι του παιδιού, που έπεσε κάτω. Στη συνέχεια το ρομπότ δεν σταμάτησε. Η εταιρεία KnightScope που ανήκε το ρομπότ με ονομασία K5 Autonomous Data Machine, ύψους 1,5 μέτρων και βάρους 136 κιλών, ζήτησε συγνώμη για το περιστατικό, έθεσε εκτός υπηρεσίας όλα τα ρομπότ ασφαλείας και εξέδωσε ανακοίνωση για το περιστατικό λέγοντας ότι το παιδί μπήκε μπροστά από το μηχάνημα. Το μηχάνημα δεν σταμάτησε και το παιδί έπεσε στο έδαφος. Οι αισθητήρες δεν αντιλήφθηκαν δόνηση και οι κινητήρες δεν σταμάτησαν, όπως θα έκαναν αν συναντούσαν εμπόδιο.

Σημειώνεται ότι πρόσφατα είδε το φως της δημοσιότητας σειρά τριών περιστατικών με τροχαία στα οποία εμπλέκονταν οχήματα της Tesla Motors που είχαν ενεργοποιημένη τη λειτουργία αυτόνομης κίνησης, εκ των οποίων μάλιστα το ένα είχε αποτέλεσμα τον θάνατο του οδηγού. Επίσης την προηγούμενη χρονιά, η αστυνομία του Ντάλας χρησιμοποίησε ένα παγιδευμένο ρομπότ εξουδετέρωσης εκρηκτικών, με εκρηκτικό μηχανισμό στον βραχίονα, για να σκοτώσει έναν ελεύθερο σκοπευτή, μετά τον θάνατο πέντε αστυνομικών και τον

⁷⁴ John Markoff, «Scientists Worry Machines May Outsmart Man», Ιούλιος 2009

τραυματισμό επτά άλλων από πυροβολισμούς. Θεωρείται πως επρόκειτο για την πρώτη θανατηφόρα χρήση «οπλισμένου» ρομπότ από την αστυνομία στις ΗΠΑ. Αυτή η κατασταλτική χρήση τους από την αστυνομία έχει προβληματίσει τους πολίτες, όπως και τα παραπάνω περιστατικά ατυχημάτων.⁷⁵

Προβληματισμούς σχετικά με το πόσο παρεμβατική μπορεί να γίνει η τεχνολογία όσον αφορά στην καθημερινή ζωή και τα δικαιώματα των πολιτών εγείρει αναπόφευκτα το Faception, πρόγραμμα «πρόγνωσης προσωπικότητας» (personality prediction) από ισραηλινή εταιρεία. Όπως λέγεται, το πρόγραμμα μπορεί να ξεχωρίσει περίπου 20 διαφορετικές ομάδες προσωπικοτήτων, από πρωταθλητές παίκτες του πόκερ (γνωστοί για τον έλεγχο που μπορούν να ασκούν στα πρόσωπά τους) μέχρι υπόπτους για εγκλήματα. Σύμφωνα με τον Σάι Γκιλμπόα, διευθύνοντα σύμβουλο της Faception, είναι δυνατή η ταυτοποίηση με υψηλή ακρίβεια των «συστατικών» προσωπικότητας, συμπεριφοράς κλπ, έτσι ώστε να μπορεί να δημιουργείται ένα προφίλ, ακόμα και σε ένα πλήθος ατόμων. «*Ας πούμε ότι σε μια πόλη επιβίβασης υπάρχουν πολλοί άνθρωποι που μπορούν δυνάμει να είναι τρομοκράτες ή βίαιο κοινό, οπότε αυτό είναι πολύ σημαντικό για τη δημόσια ασφάλεια*», είπε.

Η εταιρεία δεν αποκαλύπτει πώς λειτουργεί ο αλγόριθμος, πέραν του ότι με κάποιον τρόπο συνδέει γενετικές πληροφορίες με εκφράσεις προσώπου. Παρόλα αυτά, η Faception επιμένει πως δεν ενδιαφέρεται να διατηρεί αποθηκευμένα τα δεδομένα που συλλέγονται, και ότι δεν προκύπτουν φυλετικές διακρίσεις. Ωστόσο, ειδικοί του χώρου της αντιτρομοκρατίας υποστηρίζουν ότι η εταιρεία πρέπει να αυξήσει το ποσοστό επιτυχίας του προγράμματός της (είναι στο 86%) για να είναι χρήσιμο σε αεροδρόμια, ενώ αναμένονται αντιδράσεις από οργανώσεις προστασίας δικαιωμάτων του πολίτη.⁷⁶

Επίσης, δύο νέοι αλγόριθμοι κατασκευάστηκαν με σκοπό να διαβάζουν τα πρόσωπα. Ο ένας δημιουργήθηκε από ερευνητές του Πανεπιστημίου Στάνφορντ με επικεφαλής τον Μάικλ Κοζίνσκι και έχει τη δυνατότητα να ξεχωρίζει τους ομοφυλόφιλους από τους ετεροφυλόφιλους, εξετάζοντας τα χαρακτηριστικά και τις εκφράσεις των προσώπων. Τα πειράματα έδειξαν ότι ο αλγόριθμος μπορεί να αναλύσει τα πρόσωπα σε φωτογραφίες και να διακρίνει ανάμεσα σε «ομοφυλόφιλους» και «ετερόφυλους» με ποσοστό επιτυχίας 81% στους άνδρες και 74% στις γυναίκες. Όταν μάλιστα το λογισμικό «διάβασε» όχι μία αλλά πέντε φωτογραφίες του ίδιου προσώπου, τότε τα ποσοστά επιτυχίας του ανέβηκαν σε 91% και 83%. Συγκριτικά, οι άνθρωποι, που μελέτησαν τα ίδια πρόσωπα στις φωτογραφίες, είχαν αντίστοιχα ποσοστά επιτυχίας 61% και 54%. Σύμφωνα με τους ερευνητές, «*τα πρόσωπα περιέχουν πολύ περισσότερες πληροφορίες σχετικά με το σεξουαλικό προσανατολισμό από όσες είναι δυνατό να αντιληφθεί και να ερμηνεύσει ο ανθρώπινος εγκέφαλος*». Ο άλλος αλγόριθμος κατασκευάστηκε από ερευνητές του Πανεπιστημίου του Κέμπριτζ με επικεφαλής τον Αμαρτζότ Σινγκ και έχει τη δυνατότητα να αναγνωρίζει ένα πρόσωπο, ακόμα κι αν φοράει μάσκα. Το σύστημα αυτό έχει ποσοστό επιτυχίας 77% για όσους φοράνε μαντήλι, 69% για όσους φοράνε μαντήλι και καπέλο και 55% για όσους φοράνε μαντήλι, καπέλο και μαύρα γυαλιά.⁷⁷

Το σύστημα, που βασίζεται στην ανάλυση 14 χαρακτηριστικών-κλειδιών του προσώπου, θα μπορούσε να βρει πολλές χρήσεις, από τον εντοπισμό εγκληματιών έως ανεπιθύμητων διαδηλωτών. Αρκεί ο αλγόριθμος να «διαβάσει» έστω ένα μέρος από αυτά τα 14 χαρακτηριστικά, ιδίως γύρω από τα μάτια και το στόμα, για να μαντέψει την ταυτότητα του ανθρώπου. «*Κατ' ουσία το σύστημα μπορεί να δει διαμέσου της μάσκας*», είπε ο Σινγκ, ο

⁷⁵ James Vincent, «*Mall security bot knocks down toddler, breaks Asimov's first law of robotics*»

Ιούλιος 2016

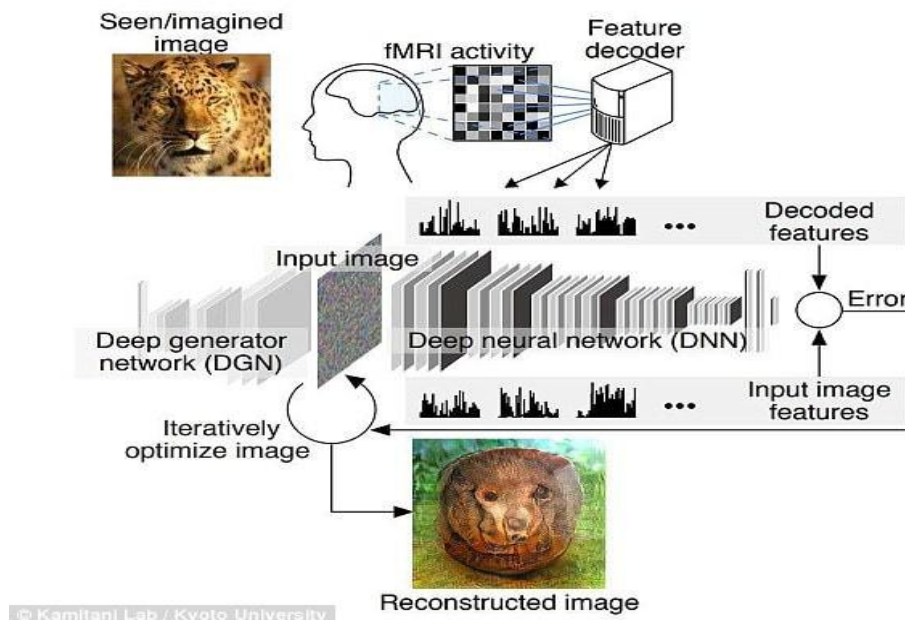
⁷⁶ Naftemporiki.gr, «*Πρόγραμμα «πρόγνωσης προσωπικότητας» για τον εντοπισμό κακοποιών*», Ιούλιος 2016

⁷⁷ ΕΛΕΥΘΕΡΙΑ online, «*Αλγόριθμοι τεχνητής νοημοσύνης ξεχωρίζουν τους "γκέι" από τους "στρέιτ"*», Σεπτέμβριος 2017

οποίος έκανε γνωστό ότι οι έρευνες θα επεκταθούν για να μελετηθεί κατά πόσο το σύστημα μπορεί να γίνει ικανό να αναγνωρίσει κάποιον που φορά άκαμπτη πλαστική μάσκα. Ο ίδιος παραδέχθηκε ότι το σύστημα «κατά κάποιο τρόπο παραβιάζει την ιδιωτικότητα των ανθρώπων» και μπορεί να αξιοποιηθεί δεόντως από αυταρχικές κυβερνήσεις.

Όλο και περισσότερες υπηρεσίες ασφαλείας ανά τον κόσμο καταφεύγουν σε λογισμικά αυτόματης αναγνώρισης προσώπου τόσο μέσα από φωτογραφίες όσο και από βίντεο. Προς το παρόν πάντως, κανένα σύστημα διεθνώς (από όσα τουλάχιστον είναι γνωστά) δεν θεωρείται τέλειο.⁷⁸

Οι Ιάπωνες επιστήμονες όμως, πήγαν ακόμα πιο μακριά. Δημιούργησαν ένα ανατριχιαστικό μηχάνημα που μπορεί να κοιτάζει στο μυαλό με απίστευτη ακρίβεια. Το AI μελετά ηλεκτρικά σήματα στον εγκέφαλο για να επεξεργαστεί ακριβώς ποιες εικόνες κοιτάζει κάποιος, ή ακόμα και σκέφτεται. Οι ειδικοί χρησιμοποίησαν ένα δίκτυο νευρώνων για την πρόβλεψη εικόνων βασισμένων σε εγκεφαλικά σήματα σε σαρώσεις. Η τεχνική στοχεύει να χρησιμοποιηθεί μια μέρα για τη δημιουργία βίντεο ονειροπολήσεων ή και κατά της ιδιωτικότητας των σκέψεων. Εξαρτάται από το ποιος και γιατί, θα την έχει στην κατοχή του.



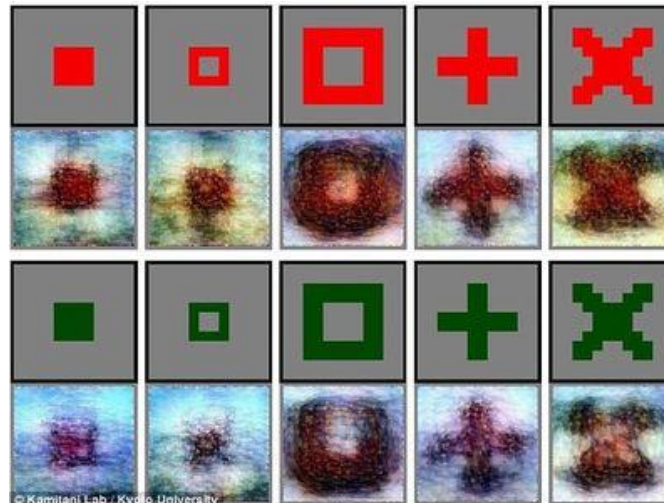
Εικόνα 8:

Ερευνητές στην Ιαπωνία δείχνουν τον τρόπο αποκωδικοποίησης της σκέψης. Επισκόπηση της αναπαράστασης βαθιάς εικόνας. Οι τιμές των εικονοστοιχείων της εικόνας εισόδου βελτιστοποιούνται έτσι ώστε τα χαρακτηριστικά DNNτης εικόνας να είναι παρόμοια με αυτά που αποκωδικοποιούνται από τη δραστηριότητα fMRI. Ένα δίκτυο βαθιάς γεννήτριας (DGN) συνδυάζεται προαιρετικά με το DNN για την παραγωγή φυσικών εικόνων, στις οποίες πραγματοποιείται βελτιστοποίηση στο χώρο εισόδου του DGN. Πηγή: bioRxiv (2017). DOI: 10.1101/240317

Το μηχάνημα αυτό δημιουργήθηκε από ερευνητές του εργαστηρίου Kamitani στο Πανεπιστήμιο του Κιότο, με επικεφαλής τον καθηγητή Yukiyasu Kamitani. Οι ειδικοί χρησιμοποίησαν ένα νευρωνικό δίκτυο για να δημιουργήσουν εικόνες βασισμένες σε πληροφορίες που ελήφθησαν από σαρώσεις fMRI, οι οποίες ανιχνεύουν αλλαγές στη ροή του

⁷⁸ ΑΠΕ-ΜΠΕ, «Αλγόριθμοι τεχνητής νοημοσύνης ξεχωρίζουν τους "γκέι" από τους "στρέιτ"», Σεπτέμβριος 2017

αίματος για να αναλύσουν την ηλεκτρική δραστηριότητα. Χρησιμοποιώντας αυτά τα δεδομένα, το μηχάνημα ήταν σε θέση να αναπαράγει κουκουβάγιες, αεροσκάφη, παράθυρα βιτρό και κόκκινα ταχυδρομικά κουτιά που τρεις εθελοντές είδαν στις φωτογραφίες. Παρήγαγε επίσης και διάφορες εικόνες αντικειμένων συμπεριλαμβανομένων τετραγώνων, σταυρών, χρυσόψαρων, κύκνων, λεοπαρδάλεων και μπάλες μπόουλινγκ που απλώς σκέφτηκαν οι συμμετέχοντες.



Εικόνα 9:

Απεικόνιση της σκέψης των εθελοντών, από Τεχνητή Νοημοσύνη.

Πηγή: Kamitani Lab / Kyoto University

Αν και η ακρίβεια ποικίλλει από άτομο σε άτομο, η ανακάλυψη ανοίγει ένα «μοναδικό παράθυρο στον εσωτερικό μας κόσμο», σύμφωνα με την ομάδα του Κιότο. Η τεχνική θα μπορούσε θεωρητικά να χρησιμοποιηθεί για να δημιουργήσει βίντεο με πλάνα των ονειροπολήσεων, μνημών και άλλων νοητικών εικόνων. Θα μπορούσε επίσης να βοηθήσει τους ασθενείς σε μόνιμες φυτικές καταστάσεις να επικοινωνούν με τους αγαπημένους τους, λένε οι επιστήμονες, προσπαθώντας να προβάλλουν μια θετική πλευρά, ίσως για να περάσει το επίτευγμά τους το τεστ ηθικής.⁷⁹

Από την άλλη, σε πρόσφατη έρευνά του ο Τζέισον Φάρμαν, επικεφαλής των οικονομικών συμβούλων του πρώην Αμερικανού προέδρου Μπαράκ Ομπάμα, προειδοποιούσε ότι η ρομποτική έχει στερήσει την αγορά από χιλιάδες εργατικά χέρια και έχει συμβάλει στην καθήλωση των μισθών, ενώ από την άλλη πλευρά έχει εκτινάξει στα ύψη τα κέρδη των επιχειρήσεων που υιοθετούν τις νέες τεχνολογίες των αυτοματισμών. Κάποιοι ειδικοί θεωρούν ότι βρισκόμαστε ενώπιον μιας νέας εποχής, χαρακτηριστικό της οποίας είναι η επανάσταση της ρομποτικής. Καθησυχάζουν όμως τους ανησυχούντες για τις κοινωνικές επιπτώσεις της επανάστασης αυτής, θυμίζοντας ότι τον 19ο αιώνα το 60% του εργατικού δυναμικού των ΗΠΑ απασχολούνταν στον αγροτικό τομέα, ενώ στις αρχές του 21ου αιώνα το ποσοστό έχει περιοριστεί στο 2% δίχως να αντιμετωπίζει η χώρα πρόβλημα απασχόλησης. «Κάτι ανάλογο θα συμβεί και τώρα. Αλλά είναι πολύ δύσκολο να προβλέψει κάποιος πώς και από πού θα προκύψουν οι νέες θέσεις εργασίας για να αντικαταστήσουν αυτές που θα κλέψουν τα ρομπότ» παρατηρεί ο Μάικλ Τσούι της εταιρείας διαχείρισης ανθρώπινου δυναμικού McKinsey.

⁷⁹ Το Φανάρι, «Τεχνητή νοημοσύνη διαβάζει το μυαλό σας και αναπαράγει τις εικόνες», Ιούλιος 2018

Ειδικοί προειδοποιούν, ότι οι σύγχρονες κοινωνίες είναι εντελώς απροετοίμαστες για να δεχθούν και να ανεχθούν την επερχόμενη επανάσταση. Άλλοι επιχειρούν να σταθμίσουν τα υπέρ και τα κατά του δίχως αμφιβολία αναπότρεπτου ρομποτικού μέλλοντός μας. Το ερώτημα γιατί είναι αναπότρεπτη η ρομποτοποίηση της καθημερινής ζωής, απαντάται εύκολα μόλις αναλογιστεί κανείς τα τεράστια ποσά που επενδύονται τα τελευταία χρόνια για την ανάπτυξη της τεχνητής νοημοσύνης. Παρεμπιπτόντως θα έπρεπε να επισημανθεί ότι πρωταγωνιστές τη φορά αυτή στις επενδύσεις του μέλλοντός μας, αναδεικνύονται αμερικανικές και κινεζικές εταιρείες εις βάρος των ιαπωνικών και των ευρωπαϊκών αλλά και των γερμανικών που κατά τεκμήριο είναι πιο εύρωστες οικονομικά. Αρκεί να αναφερθεί ότι οι επενδύσεις στον τομέα αυτόν στις ΗΠΑ υπερδιπλασιάστηκαν το 2015 συγκριτικά με το 2014 και έφθασαν στα 587 εκατ. δολάρια, σύμφωνα με την εταιρεία ερευνών CB Insights. Παγκοσμίως το 2015, οι επενδύσεις στον κλάδο της ρομποτικής αυξήθηκαν κατά 17% και έφθασαν στα 71 δισ. δολάρια, σύμφωνα με τον όμιλο ερευνών IDC, ο οποίος προβλέπει ότι το έτος 2019 θα έχουν σχεδόν διπλασιαστεί οι επενδύσεις για να φθάσουν στα 135 δισ. δολάρια.⁸⁰

5.3. Επαγγελματικές Συνέπειες

Οι επιστήμονες του MIT και οι οικονομολόγοι του Πανεπιστημίου της Οξφόρδης εκτιμούν ότι έως το 2033, θα καταργηθούν 67 εκατομμύρια θέσεις εργασίας στις ΗΠΑ εξαιτίας της τεχνητής νοημοσύνης. Αντίστοιχα, η Τράπεζα της Αγγλίας προβλέπει την κατάργηση 15 εκατομμυρίων θέσεων εργασίας στο Ηνωμένο Βασίλειο μέχρι το 2035, ενώ ο ΟΟΣΑ προβλέπει μείωση του παγκόσμιου εργατικού δυναμικού κατά 9% τα επόμενα χρόνια. Στη Γαλλία, το Συμβούλιο Ανάλυσης της Απασχόλησης εκτιμά ότι οι απειλούμενες θέσεις εργασίας είναι 11%, ποσοστό που αντιπροσωπεύει περίπου 2 εκατομμύρια εργαζόμενους.⁸¹

Συνέπειες αυτής της γοργά εξελισσόμενης τεχνητής νοημοσύνης παρατηρούνται στις δημόσιες μεταφορές, στα εμπορικά κέντρα, στη γεωργία, αλλά και στον τραπεζικό τομέα. Αυτόνομα λεωφορεία, μετρό χωρίς οδηγούς, καταστήματα χωρίς υπαλλήλους, αυτοματοποίηση της γεωργικής και κτηνοτροφικής παραγωγής, καθώς και τραπεζικών υπηρεσιών. Χαρακτηριστικό παράδειγμα είναι το λογισμικό Amelia που χρησιμοποιείται στη Σκανδιναβία, και διαχειρίζεται 10.000 λογαριασμούς, βοηθώντας έναν μόνο υπάλληλο, ενώ στη Γαλλία η αντιστοιχία είναι 500 λογαριασμοί για κάθε υπάλληλο. Άλλο ένα επίτευγμα της τεχνητής νοημοσύνης, είναι τα πρώτα μυθιστορήματα που δημιουργήθηκαν με την τεχνική αυτή, στην Ιαπωνία. Ακόμα και εάν τα υπολογιστικά συστήματα είναι ικανά να πάρουν την θέση ενός συγγραφέα, υστερούν στην απόκτηση της ευαισθησίας και των συναισθημάτων που έχει ο δεύτερος κατά την διάρκεια της συγγραφής.

Στον τομέα της ιατρικής, και ειδικότερα όσον αφορά τις ιατρικές διαγνώσεις, η τεχνητή νοημοσύνη έχει συμβάλλει τα μέγιστα. Αυτό φυσικά δεν σημαίνει απαραίτητα ότι έχει μόνο θετικές επιπτώσεις, αφού προς το παρόν τουλάχιστον υστερεί στον τομέα της ενημέρωσης και της θεραπείας του ασθενή. Όσο αφορά τον τομέα της δικαιοσύνης, η IBM έχει εφεύρει το σύστημα Ross. Με την χρήση αυτού του συστήματος, η ανάλυση χιλιάδων εγγράφων μπορεί να γίνει μέσα σε πολύ μικρό χρονικό διάστημα και να δοθούν απαντήσεις σε όλες τις υποθέσεις. Βέβαια, στις υποθέσεις δικαιοσύνης υπάρχουν και ηθικά θέματα, που είναι δύσκολο να αντιμετωπιστούν μηχανικά. Γι αυτούς τους λόγους, δεν αναμένεται να γίνει πολύ σύντομα η αντικατάσταση των δικηγόρων και των δικαστικών από τα ρομπότ.

⁸⁰ Καψύλης Αλέξανδρος, «Η τεχνητή νοημοσύνη σύμμαχος και ανταγωνιστής του ανθρώπου» Μάιος 2016

⁸¹ Κωνσταντίνος Ζοπουνίδης, Μιχάλης Δούμπος, «Τεχνητή νοημοσύνη και επαγγέλματα», Μάρτιος 2018

Όλο και πληθαίνουν οι προβλέψεις των ειδικών που αναφέρουν ότι σύντομα πολλές ειδικότητες θα εξαφανιστούν από την ανάπτυξη των μηχανών με τεχνητή νοημοσύνη. Μέχρι στιγμής, οι ρομποτικές μηχανές αναλαμβάνουν με ραγδαίους ρυθμούς χειρωνακτικά επαγγέλματα, εκτελώντας εργασίες επαναλαμβανόμενης κίνησης. Στη συνέχεια, όμως, εξυπνότερες συσκευές, θα θέσουν σε κίνδυνο επαγγέλματα που απαιτούν περισσότερες δεξιότητες, αναφέρει σε δημοσίευσμά του το CNN. Ο ρυθμός των εξελίξεων μάλιστα, θα επιταχυνθεί, καθώς αναμένεται ότι οι συσκευές που θα είναι διασυνδεδεμένες με το διαδίκτυο θα διπλασιαστούν, φτάνοντας τα 50 δισεκατομμύρια μέχρι το 2020. Το ίδιο έτος, περίπου οι μισές θέσεις εργασίας στις ΗΠΑ θα εξαφανιστούν από τους υπολογιστές και την τεχνητή νοημοσύνη, ενώ ένα επιπλέον 20% θα αντιμετωπίσει την ίδια απειλή, σύμφωνα με έρευνα του Πανεπιστημίου της Οξφόρδης. Εργασίες που στο πρόσφατο παρελθόν έμοιαζαν να μην αντιμετωπίζουν αυτόν τον κίνδυνο, βρίσκονται πλέον στον ίδιο παρονομαστή: αναλυτές δεδομένων, τραπεζίτες, ακόμη και χρηματιστές θα απειληθούν από τα ρομπότ.

Το κόστος ρομπότ και υπολογιστών συνεχώς μειώνεται, με αποτέλεσμα να αποτελούν ελκυστική επένδυση για τους επιχειρηματίες. Και όπως αναφέρει η έκθεση της Bank of America, διαρκώς γίνεται πιο εύκολη η χρήση και η εργασιακή αξιοποίηση των ρομπότ. Αναπτύσσουν ικανότητες όπως αυτή της μάθησης, αλλά και της αναγνώρισης φωνής και προσώπου, γεγονός που τα κάνει ικανά να διαχειριστούν επαγγελματικά πόστα που απαιτούν επαφή με πελάτες. Οι χώρες που μπορούν να υιοθετήσουν αυτές τις νέες τεχνολογίες θα λάβουν άμεσα μία σημαντική ώθηση στους τομείς του χαμηλού εργασιακού κόστους και της υψηλής παραγωγικότητας. Αντίστροφοι, βέβαια, θα είναι οι αριθμοί της απασχόλησης και του εργατικού δυναμικού. Σε αυτήν την τεχνολογική κούρσα προπορεύεται η Ιαπωνία, όπου ήδη αντιστοιχούν 1.520 ρομπότ ανά 10.000 εργαζομένων στα εργοστάσια κατασκευής αυτοκινήτων, νούμερο που απέχει πολύ από τον παγκόσμιο μέσο όρο (66 ρομπότ ανά 10.000). Παρόμοιο δρόμο έχει χαράξει και η Κίνα, η οποία αποτελεί τον μεγαλύτερο αγοραστή ρομπότ παγκοσμίως, καλύπτοντας το 25% της παγκόσμιας ζήτησης.

Η επερχόμενη επανάσταση θα μεταβάλει δραματικά την παγκόσμια οικονομία αλλά και θα αυξήσει ραγδαία την οικονομική ανισότητα. Οι θέσεις εργασίας που χάνονται ή κινδυνεύουν άμεσα με εξαφάνιση έχουν χαμηλότερες αποδοχές και μικρές απαιτήσεις σε δεξιότητες, σύμφωνα με τους ερευνητές. Εργαζόμενοι με πολλές γνώσεις και «πλούσιο» βιογραφικό ενδέχεται να χρειαστεί να καλύψουν θέσεις εργασίας με σχετικά απλοϊκό αντικείμενο, το οποίο όμως δεν μπορούν να εκτελέσουν μηχανές.

Σύμφωνα με τον Τάσο Οικονόμου, Ρεπόρτερ Τεχνολογίας, οι ειδικότητες που θα εξαφανιστούν γιατί θα εκτελούνται από ρομπότ, είναι αυτές των διοικητικών υπάλληλων, των αναλυτών δεδομένων, των ξεναγών, των αρτοποιών, των κρεοπωλών, των ασφαλιστών, των πωλητών σε εμπορικά καταστήματα, των εφοριακών, των πωλητών μέσω τηλεφώνου, των λογιστών, των ψαράδων, των οδηγών μέσω μεταφοράς, αλλά και αυτών που εκτελούν χειρωνακτικές εργασίες. Οι εργασίες που θα παραμείνουν στο ανθρώπινο πεδίο και δεν φαίνεται να κινδυνεύουν είναι εκείνες όπου το συναίσθημα, η δημιουργικότητα, η κοινωνική συναναστροφή, το ένστικτο είναι προαπαιτούμενα. Ανάμεσα σε αυτές είναι οι κοινωνικοί λειτουργοί, οι εργαζόμενοι στο χώρο της υγείας, οι ερευνητές, οι αστυνομικοί, οι εκπαιδευτικοί, κ.α.⁸²

«Το κύμα του ψηφιακού μετασχηματισμού και αυτοματισμού θα συνεχίσει να οδηγεί την ρομποτική έκρηξη προς τα εμπρός μέχρι το 2018», λέει ο Τζο Gemma, ο Πρόεδρος της Διεθνούς Ομοσπονδίας της Ρομποτικής. «Οι επαναστατικές εξελίξεις στον τομέα της πληροφορικής που συνδέονται με όλες τις πτυχές του Διαδικτύου, και των νέων δικτυακών υπηρεσιών αλλάζουν ριζικά τις βιομηχανίες παραγωγής. Μηχανήματα, logistics και εργοστάσια παραγωγής συγχωνεύονται σε ολοκληρωμένα κυβερνοφυσικά συστήματα. Ο

⁸² Τάσος Οικονόμου, «Ασφαλιστικές εισφορές για τα ρομπότ προβλέπει Ευρωπαϊκή πρόταση», Ιούνιος 2016

στόχος είναι η δημιουργία έξυπνων εργοστασίων που θα είναι πιο ευέλικτα, πιο οικονομικά, πιο αποδοτικά και πιο παραγωγικά.

Ο συγγραφέας Μάρτιν Φόρντ, σε συνέντευξή του στο πρακτορείο Bloomberg, ισχυρίστηκε όχι μόνο ότι είναι αναπόφευκτη η παρουσία των ρομπότ στον χώρο εργασίας αλλά και ότι θα έχουν πρωταγωνιστικό ρόλο. Γι' αυτό τον λόγο θεωρεί ότι οι εργαζόμενοι θα πρέπει να συμβιβαστούν με τα νέα δεδομένα χωρίς φόβο, και να προσαρμοστούν εκμεταλλευόμενοι τις νέες καταστάσεις σύμφωνα με το συμφέρον τους.

Από την άλλη, επιστήμονες και ειδικοί ισχυρίζονται ότι επειδή τα ρομπότ ως εργαζόμενοι δεν θα έχουν αγοραστική δύναμη ώστε να ωφελούνται επιχειρήσεις από την κατανάλωση του μισθού τους, δεν θα ενισχύεται η οικονομία της κάθε χώρας. Αντίθετα οι άνθρωποι εργαζόμενοι με τον μισθό τους ενισχύουν την οικονομία.

Ένας από τους επιστήμονες που πιστεύει πως οι τεχνητές νοημόνες και τα νοήμων υπολογιστικά συστήματα – ρομπότ- απειλούν την ανθρωπότητα με ανεργία, είναι ο Μοσέ Βαρντί (επιστήμονας υπολογιστών και καθηγητής στο Rice University). Ο ίδιος θεωρεί ότι σε λιγότερο από 30 χρόνια, περισσότεροι από τον μισό πληθυσμό του πλανήτη Γη, θα μείνει χωρίς εργασία. Για αυτό τον λόγο στην AAAS (American Association for the Advancement of Science), υποστήριξε ότι δεν θα πρέπει να αποδοθεί μικρότερη αξία στην Τεχνητή Νοημοσύνη από αυτήν που πραγματικά έχει: *«Πλησιάζουμε σε μια στιγμή που οι μηχανές θα είναι ικανές να κάνουν καλύτερα από τους ανθρώπους σχεδόν οποιαδήποτε δουλειά. Νομίζω ότι η κοινωνία πρέπει να έρθει αντιμέτωπη με αυτό το ερώτημα. Εάν οι μηχανές είναι ικανές να κάνουν σχεδόν οποιαδήποτε δουλειά κάνουν οι άνθρωποι, τι θα κάνουν οι άνθρωποι;»*.

Ο Μοσέ συγκρίνοντας την απειλή μεταξύ των στρατιωτικών drones και της Τεχνητής Νοημοσύνης, θεωρεί ότι η δεύτερη είναι μία πιο ύπουλη απειλή, αφού θα μπορούσε να οδηγήσει την ανεργία στο 50% παγκοσμίως. Η μεσαία τάξη θα εξαφανιστεί τελείως και οι ανισότητες μεταξύ των κοινωνικών στρωμάτων θα αμβλυνθούν. Η επανάσταση της Τεχνητής Νοημοσύνης θα είναι μία μάχη μεταξύ του ανθρώπινου νου και της μηχανικής σκέψης. Ακόμα και στην περίπτωση που τα ρομπότ θα χρησιμοποιηθούν από τους ανθρώπους για να κάνουν υποδεέστερες εργασίες, πάλι θα μειωθούν οι εβδομαδιαίες ώρες εργασίας και κατά συνέπεια θα μειωθεί το οικογενειακό εισόδημα. Αυτό ήδη συμβαίνει στις ΗΠΑ, σύμφωνα με στοιχεία του MIT, κάτι που έχει αντίκτυπο στην ευμάρεια των πολιτών. Επίσης τονίζει ότι καμία θέση εργασίας δεν είναι «ασφαλής» σε κανέναν επαγγελματικό τομέα, δίνοντας παραδείγματα τα «ρομπότ του σεξ» και τα «ρομπότ υψηλής ακρίβειας» που χρησιμοποιούν οι εταιρίες Foxconn και Samsung.⁸³

Το εργοστάσιο της "Foxconn" στην Κίνα, για παράδειγμα, απέλυσε 60.000 εργαζομένους και τους αντικατέστησε με ρομπότ. «Περισσότερες εταιρείες είναι πιθανόν να ακολουθήσουν αυτή την τακτική. Χρησιμοποιούμε τα ρομπότ για δύσκολες εργασίες προτρέποντας τους υπαλλήλους μας να επικεντρωθούν σε άλλες ειδικότητες που συμβάλλουν σε άλλα τμήματα της παραγωγής», ανέφερε εκπρόσωπος της "Foxconn" σε δήλωσή του στο BBC.

Σύμφωνα με την έρευνα της Διεθνούς Ομοσπονδίας Ρομποτικής, η εξαγορά ρομπότ από τη Κίνα είχε άνοδο 17% το 2015, φτάνοντας τις 68.000 μονάδες, αριθμός μεγαλύτερος από το σύνολο της Ευρωπαϊκής αγοράς. Μάλιστα, η κινεζική εταιρεία κατασκευής οικιακών συσκευών "Midea" θέλει να επεκταθεί και στην Ευρώπη. Έχουν ήδη καταθέσει πρόταση εξαγοράς της γερμανικής εταιρείας κατασκευής ρομπότ "Kuka".

Η ζήτηση των ανθρωποειδών αυξάνεται και σε άλλες ασιατικές χώρες με ταχείς ρυθμούς όπως, στη νότια Κορέα και την Ιαπωνία. Οι ΗΠΑ μάλιστα, είναι η τέταρτη μεγαλύτερη αγορά του κόσμου, που αναζητεί και αυτή νέα μέσα και καινοτομίες. Αναμένεται

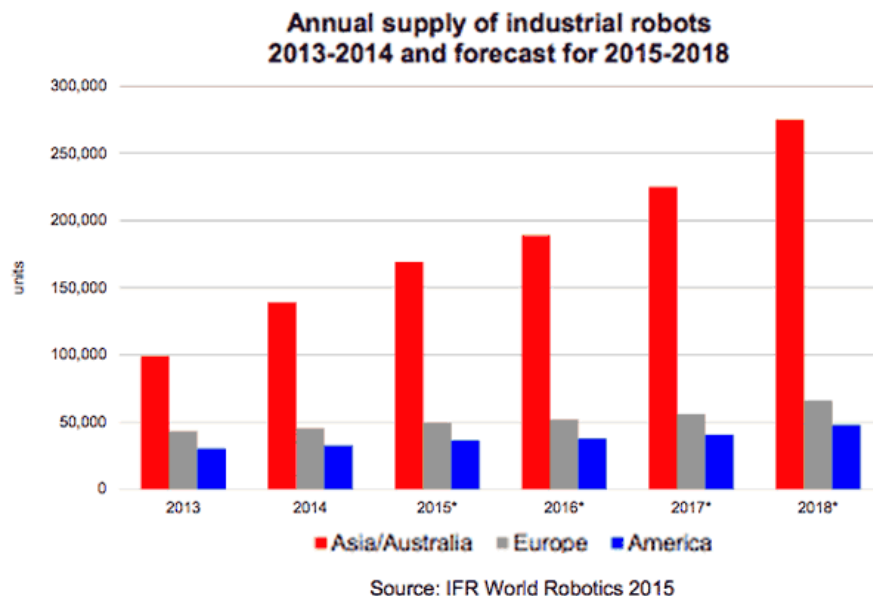
⁸³ *Nautemporiki.gr, «Δυσσιώνη πρόβλεψη: Τα ρομπότ μπορούν να αφήσουν άνεργο τον μισό πληθυσμό μέσα σε 30 χρόνια», Φεβρουάριος 2016*

ότι περίπου 2.3 εκατομμύρια ρομπότ θα αναπτυχθούν σε ορόφους εργοστασίων σε όλον τον κόσμο από το 2018, δηλαδή περισσότερο από το διπλάσιο του ποσού το 2009. Σήμερα είναι πιο ευρέως χρησιμοποιούμενα στον κλάδο των αυτοκινήτων. Αλλά η δημοτικότητά τους αυξάνεται και στη μεταλλουργία, στην πλαστική ύλη και στους τομείς της ηλεκτρονικής.

Η χρήση τους για βρώμικες, επικίνδυνες και δύσκολες εργασίες, όπως προ είπαμε, κάνει τη ζωή μας ευκολότερη και πιο ασφαλή. Σε δειγματοληπτική έρευνα που πραγματοποίησε η IFR (international federation of robotics – Διεθνής Ομοσπονδία Ρομποτικής) σε 7.000 εργαζόμενους σε επτά χώρες, σχεδόν το 70% δήλωσαν ότι η ρομποτική και η αυτοματοποίηση στην παραγωγική διαδικασία θα τους δώσουν την ευκαιρία να διεκδικήσουν πιο εξειδικευμένες και καλύτερα αμειβόμενες θέσεις εργασίας.

Νέα μελέτη της IFR με θέμα «Ρομπότ και η θέση εργασίας του μέλλοντος» παρουσιάζει αισιόδοξη εικόνα, σύμφωνα με την οποία στην επόμενη δεκαετία η ποιότητα εργασίας και οι αμοιβές θα βελτιωθούν, ενώ νέου τύπου θέσεις εργασίας θα δημιουργηθούν. Επιπλέον δημιουργήθηκε ένα καινούριο τεχνητό ρομποτικό σύστημα που προστατεύει όχι μόνο τα ρομπότ, αλλά και τους ανθρώπους εργαζόμενους, από ατυχήματα.

Η IFR επισημαίνει ότι, παρά τους διάχυτους φόβους στα μέσα ενημέρωσης και στην κοινή γνώμη, «δεν υπάρχουν καθόλου απτά στοιχεία που να δείχνουν ότι η τρέχουσα αυτοματοποίηση είναι διαφορετική στις επιπτώσεις της πάνω στην απασχόληση από ό,τι τα προηγούμενα κύματα τεχνολογικών αλλαγών στην πορεία των αιώνων, τα οποία όλα οδήγησαν σε ένα μείγμα απώλειας θέσεων εργασίας, δημιουργίας νέων θέσεων και αλλαγής του προφίλ του τρόπου της δουλειάς».⁸⁴



Εικόνα 10:

IFR 2015, Ετήσιος εφοδιασμός βιομηχανικών ρομπότ σε Ασία/Αυστραλία, Ευρώπη και Αμερική

Εκτός από τον «Φίλιπ Ντικ», το εκφραστικό ανθρωποειδές ρομπότ που βοηθάει τους γιατρούς στις διαγνώσεις και στο οποίο θα αναφερθούμε διεξοδικά σε επόμενο υποκεφάλαιο, Γερμανοί ερευνητές του Πανεπιστημίου Leibniz στο Αννόβερο, ανέπτυξαν ένα τεχνητό ρομποτικό νευρικό σύστημα που θα «διδάσκει στα ρομπότ να νιώθουν πόνο» και να

⁸⁴ ΑΜΠΕ, «Πάνω από 3 εκατομμύρια βιομηχανικά ρομπότ θα λειτουργούν το 2020», Ιολυνιος 2018

αντιδρούν γρήγορα για να αποφεύγουν ζημιά στα εξαρτήματά τους. Οι ερευνητές παρουσίασαν τη δουλειά τους τον Μάιο του 2016 στην IEEE International Conference on Robotics and Automation (ICRA) στη Στοκχόλμη της Σουηδίας. Μιλώντας στο IEEE Spectrum, ο Γιόχαν Κουν, ένας εξ αυτών, εξηγεί ότι ο πόνος στην ουσία είναι ένα σύστημα προστασίας, αφού «Όταν αποφεύγουμε την πηγή του πόνου, μας βοηθά να μην πληγωθούμε». Ο Κουν συνεργάστηκε στο πρόγραμμα με τον καθηγητή Σάμι Χανταντίν, έναν από τους κορυφαίους ειδικούς του κόσμου στη φυσική αλληλεπίδραση ανθρώπων και ρομπότ, και υποστηρίζει πως το σύστημα αυτό θα προστατεύει όχι μόνο τα ρομπότ από βλάβες, αλλά και τους ανθρώπους που θα εργάζονται σε μικρή απόσταση από τα ρομπότ, μειώνοντας τις πιθανότητες, προβλήματα που δεν έχουν εντοπιστεί να μπορούν να προκαλέσουν ατυχήματα.

Το όλο σύστημα χρησιμοποιεί ένα «νευρικό μοντέλο ρομποτικού ιστού που εμπνέεται από τη δομή του ανθρώπινου δέρματος» για να «αποφασίζει πόσο πόνο θα έπρεπε να νιώθει το ρομπότ, σε μια δεδομένη ποσότητα δύναμης. Όπως οι άνθρωποι νευρώνες, το μοντέλο μεταδίδει πληροφορίες πόνου σε επανειλημμένες «εξάρσεις» αν η δύναμη που ασκείται ξεπεράσει ένα όριο, και το σχετικό «χειριστήριο πόνου» αντιδρά, αφού πρώτα καταχωρήσει τις πληροφορίες ως ελαφρύ, μέσο ή μεγάλο πόνο. Το ρομπότ «νιώθει» άσχημα και αποσύρεται λίγο μόλις παύσει το γεγονός που προκαλεί τον πόνο, για να επιστρέψει μετά σε περίπτωση μικρού πόνου. Σε περίπτωση μέσου, αποσύρεται πιο μακριά, ενώ σε περίπτωση μεγάλου λαμβάνονται έκτακτα μέτρα, καθώς θεωρείται πως απαιτείται βοήθεια. Όπως αναφέρεται στο δημοσίευμα του IEEE Spectrum, το paper αυτό αποτελεί μόνο το πρώτο βήμα προς την κατεύθυνση ενός rain-based reflex controller, το οποίο μπορεί να ακούγεται κάπως «απάνθρωπο» ως ιδέα, αλλά αναμένεται να φανεί ιδιαίτερα χρήσιμο για τα ρομπότ.⁸⁵

Η συνεχής τεχνολογική εξέλιξη προετοιμάζει την ανθρωπότητα για την επικείμενη ανάπτυξη της υψηλής τεχνητής νοημοσύνης από τα ρομπότ. Όπως γράφει η Deutsche Welle, οι ηλεκτρονικοί υπολογιστές γίνονται ολοένα εξυπνότεροι, το βασικότερο όμως όλων είναι ότι πλέον έχουν αναπτύξει την ικανότητα να μαθαίνουν και να επεξεργάζονται μόνοι τους νέες πληροφορίες. Ο καιρός κατά τον οποίο τα ρομπότ μπορούσαν μόνο να εκτελούν προγραμματισμένες εντολές έχει παρέλθει ανεπιστρεπτή. Σήμερα χάρη σε νέα αλγοριθμικά μοντέλα μπορούν μέσα σε ελάχιστο χρόνο να επεξεργαστούν τεράστιες ποσότητες σύνθετων δεδομένων. Έτσι, κάνουμε τη σκέψη ότι όλοι οι κλάδοι της ανθρώπινης εργασίας κινδυνεύουν να αντικατασταθούν από τις τεχνητές νοημοσύνες. Σε πόσο χρονικό διάστημα όμως θα μπορούσε να συμβεί κάτι τέτοιο, και εάν συμβεί ποιος θα είναι ο κύριος καταναλωτής;

Ένα ακόμα παράδειγμα, είναι ένα εστιατόριο με σερβιτόρες ρομπότ που άνοιξε στην επαρχία Guizhou της Κίνας. Τα πέντε ρομπότ που διαθέτει, είναι προγραμματισμένα να σερβίρουν τους πελάτες, μεταφέροντας πιάτα στα τραπέζια μετά από προκαθορισμένες παραγγελίες, κινούμενα μόνο σε συγκεκριμένους διαδρόμους και δε δέχονται παραγγελίες από τους πελάτες. Ο μέσος μισθός για τους εργαζομένους είναι περίπου 368 δολάρια το μήνα. Τα ρομπότ έχουν «ζωή» πέντε με οχτώ χρόνια. Ένα άλλο εστιατόριο στην Κίνα χρησιμοποιεί ρομπότ τόσο στην κουζίνα όσο και στο σέρβις. Σύμφωνα με τον ιδιοκτήτη αυτού του εστιατορίου καθένα κοστίζει 6.500 δολάρια. Αυτά τα ρομπότ μπορούν να κατανοούν σαράντα εκφράσεις της καθημερινότητας και μπορούν να ζητήσουν από τους πελάτες να φύγουν από τον δρόμο τους. Για πέντε ώρες εργασίας χρειάζονται δύο ώρες φόρτιση.

Σύμφωνα όμως με πολλά ιντερνετικά sites,⁸⁶ τα ρομπότ δεν κατάφεραν να αντικαταστήσουν επάξια τους ανθρώπους σερβιτόρους, γιατί συχνά έχυναν τα ποτά και τις σούπες και έριχναν κάτω τις παραγγελίες. Δύο από τα τρία εστιατόρια στη νότια Κίνα, που

⁸⁵ *Nautemporiki.gr*, «Ρομπότ που νιώθουν πόνο», Ιανουάριος 2016

⁸⁶ *enikos.gr* 12/04/2016, *eurinika.gr* 08/04/2016, *pontos-news.gr* 10/04/2016, *lifo.gr* 08/04/2016, *typosthes.gr* 13/04/2016

χρησιμοποιούσαν ρομπότ σερβιτόρους, έκλεισαν και το τρίτο δουλεύει τώρα μόνο με ανθρώπινο προσωπικό.

Η παγκόσμια οικονομία βρίσκεται σε «τροχιά σύγκρουσης» δεδομένης της τεχνολογικής προόδου, καθώς οι προσδοκίες εκατομμυρίων εργαζόμενων έρχονται σταδιακά αντιμέτωπες με τη μελλοντική αυτοματοποίηση αρκετών θέσεων εργασίας, προειδοποίησε τον Οκτώβρη του 2017, ο πρόεδρος της Παγκόσμιας Τράπεζας, Τζιμ Γιονγκ Κιμ. Μιλώντας σε συγκέντρωση στη Νέα Υόρκη, επίσης υπογράμμισε τη σημασία της επιπλέον θεσμικής προάσπισης και των επενδύσεων στη μόρφωση όσο και τις δομές υγείας, ώστε οι εργαζόμενοι να παραμένουν ανταγωνιστικοί και ασφαλείς σε ένα διαρκώς εναλλασσόμενο οικονομικό περιβάλλον.

«Καθώς στο μέλλον τα ρομπότ θα αντικαθιστούν εκατομμύρια ανειδίκευτους εργαζόμενους και υπαλλήλους χαμηλών δεξιοτήτων, θα χρειαστεί μία διαφορετική στρατηγική επενδύσεων που θα διασφαλίζει την επένδυση στον ανθρώπινο παράγοντα. Αν οι προσδοκίες των εργαζόμενων για δουλειά δεν βρουν ανταπόκριση και ανάλογες ευκαιρίες, η κοινωνική δυναμική θα μπορούσε να οδηγήσει στην ευθραυστότητα, τις συγκρούσεις και ακόμα και την βία», παρατήρησε ο Τζιμ Γιονγκ Κιμ. Πολλοί επιστήμονες και διανοητές συνεχίζουν να προειδοποιούν για τις ενδεχόμενες ανισότητες στις οποίες θα οδηγήσει η εξέλιξη της τεχνολογίας και η «μοιραία» αντικατάσταση του ανθρώπινου δυναμικού από ρομποτικές διατάξεις.⁸⁷

Το Δεκέμβρη του 2016, ο Λευκός Οίκος έδωσε στη δημοσιότητα έκθεση για την αυτοματοποίηση, την τεχνητή νοημοσύνη και την οικονομία, προειδοποιώντας: *«Ελλοχεύει κίνδυνος να παραγκωνιστούν εκατομμύρια Αμερικανοί από την εργασία και να χάσει η χώρα τη θέση της ως παγκόσμια οικονομική ηγετική δύναμη».*

Στον αποχαιρετιστήριο από το Λευκό Οίκο λόγο του, ο Μπαράκ Ομπάμα μίλησε για αγώνα συμπαράστασης στους ανθρώπους που είναι εκτεθειμένοι στην επέλαση της τεχνολογίας. *«Το επόμενο κύμα των οικονομικών ανατροπών θα προέλθει από το αδυσώπητο κύμα της αυτοματοποίησης, που καταστρέφει θέσεις εργασίας εις βάρος της μεσαίας τάξης»,* είπε, σύμφωνα με τους «New York Times».

Δεν είναι μάλλον τυχαίο ότι ο Ντ. Τραμπ, νυν πρόεδρος των ΗΠΑ, είχε σύνθημα προεκλογικό, για επιστροφή αμερικανικών επιχειρηματικών ομίλων που παράγουν εκτός ΗΠΑ, για δημιουργία έργων υποδομών αλλά και νέων εργοστασίων, ώστε να δημιουργηθούν θέσεις εργασίας. Μάλιστα, αυτά ειπώθηκαν σε περίοδο που επίσημα, στατιστικά, έχουν μειώσει την ανεργία κάτω από 5%, πλασματικά όμως, αφού γύρω στα 49 εκατομμύρια ανέργων δεν καταγράφονται.

Ο Μπιλ Γκέιτς, ιδιοκτήτης της «Microsoft» που παράγει νέες τεχνολογίες, σε συνέντευξή του, είπε: *«Την ίδια στιγμή που προωθείται ο αυτοματισμός και χάνονται θέσεις εργασίας, πρέπει να αναπτυχθούν υπηρεσίες για το καλό των ανθρώπων, και ιδιαίτερα των φτωχότερων και πιο αδύναμων, τους γηραιότερους, τα παιδιά με ιδιαίτερες ανάγκες, τα σχολεία, εκεί όπου υπάρχουν λιγότεροι για να τους υποστηρίξουν και απαιτείται ένα ιδιαίτερο ανθρώπινο γνώρισμα, η εν-συναίσθηση. Δεν μπορούμε να αγνοήσουμε ότι η αντικατάσταση ανθρώπων από ρομπότ μηδενίζει το εισόδημα από το φόρο που κατέβαλλε αυτός ο εργαζόμενος, φόρος από τον οποίο χρηματοδοτούνται σήμερα οι προαναφερόμενες υπηρεσίες. Εξάλλου, η κοινωνία δεν θα πρέπει να φοβάται την καινοτομία, θα πρέπει να ενθουσιάζεται από αυτήν και γι' αυτό θα ήταν καλύτερα να γνωρίζει ότι υπάρχει ένα ξεκάθαρο πλάνο για το τι θα συμβεί στους ανθρώπους που χάνουν τη δουλειά τους από την*

⁸⁷ Πάνος Σάκκας, «Παγκόσμια τράπεζα τα εργαζόμενα ρομπότ θα οδηγήσουν στην κοινωνική βία», Οκτώβριος 2017

έλευση των ρομπότ. Οι επιχειρήσεις δεν θα αναλάμβαναν αυτή την πρωτοβουλία, οι κυβερνήσεις όμως θα έπρεπε».⁸⁸

5.4 Θεολογικές επιπτώσεις από την εξέλιξη της υπολογιστικής νοημοσύνης.

Ο μέσος άνθρωπος αισθάνεται ολοένα λιγότερο - αντίθετα με το παρελθόν - ότι αποτελεί «τον ήρωα του μέλλοντος». Εν έτη 2018, κατακλύζεται πια από όρους (μηχανική μάθηση, γενετική μηχανική, blockchain κ.α.), που αισθάνεται ότι κατά βάση δεν τον αφορούν.

Το Southern Evangelical Seminary & Bible College (SES), κολλέγιο στις ΗΠΑ, αγόρασε ένα αυτόνομο ανθρωποειδές ρομπότ, το NAO (Aldebaran Robotics), προκειμένου να μελετήσει τις επιπτώσεις της γοργά αναπτυσσόμενης ρομποτικής τεχνολογίας, από ηθική και θεολογική σκοπιά. Τα ερωτήματα που απασχόλησαν τους επιστήμονες ήταν αν τα ρομπότ θα μπορέσουν να αντικαταστήσουν τους ανθρώπους στην φροντίδα ασθενών ή ακόμα και παιδιών και εάν η απομάκρυνση της ανθρώπινης επαφής παραβιάζει την ηθική.

Το παραπάνω κολλέγιο αναφέρει ότι: «*Η απόκτηση του ρομπότ NAO θα επιτρέψει στους σπουδαστές και το διδακτικό προσωπικό να διεξάγουν έρευνα πάνω στην ηθική των ανερχόμενων τεχνολογιών, στις οποίες συμπεριλαμβάνονται η ρομποτική, η βιονική, η βελτίωση του ανθρώπου, η τεχνητή νοημοσύνη, ο μετανθρωπισμός και η νανοτεχνολογία, καθώς και να αποκτήσουν γνώση σχετικά με τα πολλά ηθικά ερωτήματα που προκύπτουν στο πλαίσιο αυτών των μελετών*».

Σε ανάρτησή της, η ιστοσελίδα RNS (Religion News Service), αναφέρει ότι το ρομπότ στοίχισε 9300 δολάρια στο SES. Το ρομπότ αποτελεί σημαντικό κομμάτι και άλλων ερευνητικών ιδρυμάτων, όπως το MIT, το Πανεπιστήμιο του Τόκιο και το Carnegie Mellon. Ο Κέβιν Στάλεϊ (καθηγητής θεολογίας του SES), σε συνέντευξή του στο Popular Science δήλωσε ότι: «*Όταν έρθει η ώρα της συμπερίληψης ή ενσωμάτωσης ανθρωποειδών ρομπότ στην κοινωνία, η προοπτική μίας σπασμοδικού χαρακτήρα αντίδρασης από τη θρησκευτική κοινότητα είναι εξαιρετικά πιθανή, εκτός και αν αρχίσει να υπάρχει κάποιου είδους διάλογος και αρχίσουμε να εξετάζουμε το θέμα εκτενέστερα*».

Επίσης, πιστεύει ότι η χωρίς όρια χρήση της ρομποτικής τεχνολογίας δεν είναι σωστή: «*Από θρησκευτικής άποψης, θα μπορούσε να έχει να κάνει περισσότερο με το να βλέπει κανείς τα ανθρώπινα όντα ως κατ' εικόνα και καθ' ομοίωσίν του Θεού. Και όσον αφορά σε ένα ανθρωποειδές ρομπότ, κάποιοι θα μπορούσαν να το εκλάβουν ως “κακό”, εξαιτίας της “προσπάθειας” να μιμηθεί κάτι το οποίο δεν θα έπρεπε να μιμηθεί*». Ο ίδιος μας προειδοποιεί για τον θεολογικό κίνδυνο που θα προκύψει εξ αιτίας της χρήσης των ρομπότ στην φροντίδα των ανθρώπων ή την χρήση τους ως φίλους και εραστές.⁸⁹

Η τεχνολογία έως τώρα, έχει επικεντρωθεί στην επιδιόρθωση καταστραμμένων συστημάτων. Ωστόσο οι επιστήμονες ισχυρίζονται ότι η ανθρωπότητα δεν θα σταματήσει σε αυτήν την εξέλιξη, αλλά θα καταφέρει το ακατόρθωτο. Θα έρθει η στιγμή που οι σωματικές πληγές θα επουλώνονται με τη χρήση ενός μικροτσίπ, και άλλα πολλά που φαίνονται αδιανόητα να συμβούν το 2019. Όλη αυτή όμως η τεχνολογική εξέλιξη, μπορεί να έχει ως αποτέλεσμα την δημιουργία ενός νέου πολύ σοβαρού ηθικού ζητήματος. Ένα ηθικό ζήτημα όχι μόνο σχετικά με την προοπτική του εκ νέου σχεδιασμού του εαυτού μας αλλά και της οικογένειάς μας. Οι άνθρωποι που υποστηρίζουν αυτήν την ιδέα, υποστηρίζουν ότι: «εάν συνεχίσουμε προς την ανασχεδίαση του ανθρώπινου είδους τότε η ζωή του μέλλοντος θα

⁸⁸ Nomisma.com, «Ο τρόμος του Τραμπ: Τεράστια ανεργία λόγω τεχνολογικής έκρηξης», Ιανουάριος 2017

⁸⁹ Naftemporiki.gr, «Διερευνώντας τις ηθικές και θεολογικές επιπτώσεις της ρομποτικής», Φεβρουάριος 2014

αλλάξει και αυτό θα έχει ως αποτέλεσμα την απόκτηση από τον άνθρωπο “θεϊκών” δυνάμεων».⁹⁰

Ας επικεντρωθούμε όμως στο παρόν και σε μια θεωρία που ανέπτυξαν οι μετανθρωπιστές. Παρεπιπτώντος, μετανθρωπισμός είναι ένα κίνημα, μια θεωρία που θεωρεί ότι ο συνήθης άνθρωπος αποτελεί απλά ένα "στάδιο ανάπτυξης" της συνολικής ανθρώπινης οντότητας. Ακριβέστερα, ο μετανθρωπισμός πρεσβεύει ότι ο θάνατος αποτελεί απλά την μετάβαση σε επόμενο Στάδιο Πολυπλοκότητας κατά το οποίο ο άνθρωπος εξελίσσεται σε μετάνθρωπο, ο οποίος δεν θεωρείται ως μετεξελικτικό στάδιο του ανθρώπου όπως το εννοεί η Εξελικτική Θεωρία, αλλά αντίθετα αποτελεί στάδιο ανάπτυξης. Ο μετανθρωπισμός δέχεται ότι κάθε άνθρωπος αντιπροσωπεύεται από μία προσωπικότητα, ψυχή. Αυτή η "ψυχή" (αντίστοιχη με το software του Υπολογιστή) διέρχεται από το υλικό στάδιο του ανθρώπου, κατά το οποίο συγκεντρώνει πληροφορίες κάθε είδους (βιολογικές, επιστημονικές κ.α.) και στην συνέχεια περνά στα επόμενα μετανθρωπικά στάδια.

Οι μετανθρωπιστές λοιπόν πιστεύουν, πως τελικά η μεγαλύτερη αλλαγή στη διάρκεια των χιλιετιών είναι η επινόηση μιας οντότητας που θα έχει τη δυνατότητα να είναι αθάνατη. Σε προηγούμενες εποχές ήταν η ψυχή, το "πνεύμα" που έπρεπε όμως να ξεπεράσει το "σαρκίο" που κουβαλούσε. Σήμερα, όλα σχετικοποιούνται με τη δυνατότητα μεταφοράς της συνείδησης σε έναν υπολογιστή. Όσο για το σώμα, μπορεί να είναι βιονικό. Αυτό ακούγεται σαν ένα νέο είδος θεολογίας. Όπως ο Κουρτσβέλ αναφέρει στο βιβλίο του "Η Μοναδικότητα είναι κοντά (The Singularity is Near)", "*χρειαζόμαστε μια νέα θρησκεία*". Περιγράφει μια "νοημοσύνη πυριτίου" που θα προκύψει από τη στιγμή που θα ξεπεράσουμε τη βιολογική νοημοσύνη, όπως ισχυρίζεται. Αυτός είναι για τους μετανθρωπιστές ο νέος και ο ολοκληρωμένος Θεός, μια ύστατα έξυπνη οντότητα που θα ξυπνήσει, επειδή εμείς οι άνθρωποι με την υπάρχουσα νοημοσύνη μας "*έχουμε κορέσει την ύλη και την ενέργεια του σύμπαντος*".

Η Απεροπία (ελληνικό ιστολόγιο περί μετανθρωπισμού), θεωρεί ότι οι ιεράρχες των πολυθεϊστικών θρησκειών δεν θα έχουν ιδιαίτερο πρόβλημα να δεχτούν τα νέα κοσμολογικά δεδομένα, ενώ οι ιεράρχες του μονοθεϊσμού θα αντιμετωπίσουν μεγάλες δυσκολίες. Έτσι θεωρεί, ότι δεν θα αργήσει η μέρα που η συμφωνία όλων των παγκόσμιων θρησκευτικών ηγετών για θεολογική σύγκλιση και παν-θρησκεία θα ανακοινωθεί στους πιστούς όλων των θρησκειών της Γης. Ο Kurzweil, ο οποίος είναι διευθυντής μηχανικής της Google, πιστεύει ότι οι εξελίξεις στην τεχνολογία και στην ιατρική, μας ωθούν προς τη Μοναδικότητα, σε μια περίοδο βαθιάς πολιτιστικής και εξελικτικής αλλαγής στην οποία οι υπολογιστές θα αντικαταστήσουν τον ανθρώπινο εγκέφαλο και θα επιτρέπουν στους ανθρώπους, να ζήσουν για πάντα. Τοποθετεί δε αυτή την εξέλιξη στο 2045.⁹¹

Ωστόσο και στο χώρο λατρείας διαφόρων θρησκευμάτων, αρχίζουν να εμφανίζονται κάποιες καινοτομίες που έχουν να κάνουν με την εξέλιξη της τεχνολογίας, σε μια προσπάθεια να δοθεί μια θεολογική προσέγγιση σε μια μηχανή. Όσο κι αν ακούγεται παράξενο ή αντιφατικό, μια ομάδα προτεσταντών ιερέων, έφτιαξαν ένα ιερέα-ρομπότ. Στο πλαίσιο των εορτασμών για τα 500 χρόνια από τότε που ο Μαρτίνος Λούθηρος θυροκόλλησε στην πύλη μιας εκκλησίας, τις 95 «Θέσεις» του, ζητώντας δραστική μεταρρύθμιση της Ρωμαιοκαθολικής Εκκλησίας. Μια ομάδα από την Προτεσταντική Εκκλησία της Έσσης και του Νασάου, με επικεφαλής τον κληρικό Στέφαν Κρεμπς, κατασκεύασε έναν ιερέα-ρομπότ, που ακούει στο όνομα BlessU-2 «Την ευλογία μου (να έχετε) κι εσείς, επίσης». «*Θέλαμε να εξετάσει ο κόσμος κατά πόσον είναι δυνατόν να τον ευλογήσει μια μηχανή, ή χρειάζεται ένας άνθρωπος*» δήλωσε ο Κρεμπς.

⁹⁰ Δημήτριος Θεοφίλου, «Τρανσουμανισμός, Τεχνο-θεολογία, Ανθρώπινο πρόσωπο και Θεολογική νοηματοδότηση», Δεκέμβριος 2010

⁹¹ Χρυσούλα Μπουκουβάλα, «Οι μάσκες μίας απολυταρχικής εξουσίας και η μοναδικότητα», Μάιος 2017



Εικόνα 11:
Ο ιερέας-ρομπότ «BlessU-2» ευλογεί.

Ο ιερέας-ρομπότ διαθέτει οθόνη αφής, κεφάλι και δυο χέρια και δίνει την ευλογία του στα γερμανικά, αγγλικά, γαλλικά, ισπανικά ή πολωνικά, αναλόγως την επιλογή αυτού που επιθυμεί να ευλογηθεί. Επιπροσθέτως, έχει τη δυνατότητα να επιλέξει μεταξύ γυναικείας και ανδρικής φωνής. Ο ιερέας-ρομπότ υψώνει τα χέρια του (προς τον Ύψιστο), αναβοσβήνουν τα φωτάκια του, απαγγέλει μια ρήση από τη Βίβλο και λέει: «Ο Θεός σε ευλογεί και σε προστατεύει». Αν του ζητηθεί, προσφέρει και εκτύπωση των όσων είπε. «*Η όλη ιδέα είναι να προκαλέσουμε δημόσιο διάλογο*» εξήγησε ο Κρεμπς. «*Ο κόσμος απ' το δρόμο δείχνει περιέργεια, ενδιαφέρον, διασκεδάζει. Μέσα στην εκκλησία, όμως, ο κόσμος νομίζει ότι θέλουμε να αντικαταστήσουμε τους πάστορες με μηχανές*» πρόσθεσε. Για τον Κρεμπς, ένα ρομπότ ποτέ δεν μπορεί να προσφέρει ό,τι προσφέρει ο ιερωμένος στον πιστό: «*Δεν θέλουμε να ρομποτοποιήσουμε τη δουλειά μας στην εκκλησία, αλλά θέλουμε να δούμε κατά πόσον μπορούμε να δώσουμε θεολογική προοπτική σε μια μηχανή*». Πάντως, το ρομπότ BlessU-2 δεν είναι το πρώτο ρομπότ που εισέβαλε στον κόσμο των πιστών. Πριν 2 χρόνια, βουδιστικός ναός στην περιφέρεια του Πεκίνου παρουσίασε έναν μοναχό-ρομπότ που ψέλνει βουδιστικά και εξηγεί τις βασικές θέσεις του βουδισμού.⁹²

Η τεχνολογία όμως, δεν αρκείται στο να χρησιμοποιηθεί ως αντικείμενο προσηλυτισμού αλλά στο να μπορεί να χρησιμοποιηθεί και σαν λατρευτικό είδωλο. Αξίζει να σημειωθεί ότι τα έξοδα για την ίδρυση καινούριας θρησκείας μειώνονται σημαντικά με την χρήση της τεχνολογίας. Χρησιμοποιώντας τα μέσα κοινωνικής δικτύωσης, η προσέγγιση των πιστών γίνεται με ελάχιστο κόστος καθώς με ένα πάτημα κουμπιού είναι εφικτή η δημιουργία ενός νέου Θεού. Χαρακτηριστικό παράδειγμα είναι αυτό του Anthony Levandowski, μέχρι πρότινος στέλεχος της Uber και της Google. Ο ίδιος, τον Σεπτέμβριο του 2015 στην Καλιφόρνια, ίδρυσε την Εκκλησία «Way of the Future». Η εκκλησία αφιερώθηκε στην λατρεία ενός Θεού που στηρίζεται στην τεχνητή νοημοσύνη.

Σε μία συνέντευξη που έδωσε στις 15 Νοεμβρίου του 2017 στο Wired, ο Levandowski δήλωσε ότι: «*Δεν πρόκειται για έναν Θεό με την έννοια ότι ρίχνει κεραυνούς ή προκαλεί τυφώνες. Ωστόσο, υπάρχει κάτι το οποίο είναι ένα δισεκατομμύριο φορές πιο έξυπνο από τον άνθρωπο. Πώς θα το λέγατε εσείς αν όχι Θεός;*». Αυτή η νέα θρησκεία όχι μόνο αναγνωρίστηκε από την Ομοσπονδιακή κυβέρνηση των ΗΠΑ, αλλά απαλλάχτηκε και από την υποχρέωση καταβολής φόρου (από το International Revenue Service).

⁹² ΑΠΕ-ΜΠΕ, «Ο ιερέας-ρομπότ “BlessU-2” ευλογεί», Ιούνιος 2017

Η καινούρια αυτή Εκκλησία επικεντρώνεται: *«στην αναγνώριση και λατρεία μιας θεότητας η οποία βασίζεται στην τεχνητή νοημοσύνη (AI) η οποία αναπτύσσεται με τη βοήθεια υλικού πληροφορικής και προγραμμάτων»*. Η θρησκεία που δημιούργησε ο Anthony Levandowski αφορά ως επί το πλείστον επιστήμονες και εργαζομένους της τεχνητής νοημοσύνης, αλλά, επίσης, και *«κοινούς θνητούς, οι οποίοι ενδιαφέρονται για την λατρεία μιας θεότητας βασισμένης στην τεχνητή νοημοσύνη»*. Ο ίδιος έχει σκοπό να οργανώσει ένα *«κατελιέ και εκπαιδευτικά προγράμματα στην κοιλάδα του Σαν Φραντσίσκο, μέσα στο 2018»* ώστε όλοι να ενημερωθούν.

Ενώ ο Levandowski θα μπορούσε να δημιουργήσει μια επιχείρηση ή μια ομάδα σκέψης, αποφάσισε να αφιερώσει μια λατρεία στην AI. Σύμφωνα με τα λεγόμενα του, ο λόγος που το έκανε είναι για να προσεγγίσει το μεγαλύτερο ποσοστό της ανθρωπότητας και να αποτρέψει κάποιους να υποπτευθούν ότι η αγάπη που έχει στην τεχνητή νοημοσύνη κρύβει συμφέροντα. *«Ήθελα να το κάνω με έναν τρόπο που θα επιτρέπει σε όλο τον κόσμο να συμμετέχει [...]. Αυτό αφαιρεί, επίσης, από τους ανθρώπους τη δυνατότητα να πουν ότι το κάνω μόνο για τα χρήματα»*. Επίσης δηλώνει βέβαιος ότι θα έρθει η ημέρα που οι υπολογιστές θα υπερβούν τον άνθρωπο και θα τον οδηγήσουν σε μια καινούρια εποχή, που ονομάζεται *«Singularity (μοναδικότητα)»*. Ο ίδιος έχει ονομάσει αυτήν πρόοδο *«μετάβαση»* και χαρακτηριστικά αναφέρει ότι: *«Τα ανθρώπινα όντα είναι υπεύθυνα για τον πλανήτη, γιατί είμαστε πιο έξυπνοι από τα άλλα ζώα, είμαστε ικανοί να κατασκευάζουμε εργαλεία και να εφαρμόζουμε κανόνες»*.

Μελλοντικά ισχυρίζεται, ότι το Ίντερνετ θα έχει την δυνατότητα να συνδέει τα έξυπνα τηλέφωνα μεταξύ τους, όπως και το νευρικό σύστημα στον άνθρωπο, με τα κέντρα δεδομένων να είναι ο εγκέφαλος. Αυτός ο εγκέφαλος θα έχει την ικανότητα να παρακολουθεί τα πάντα, να αντιλαμβάνεται τα πάντα και να έχει γνώσεις για τα πάντα. Για αυτόν τον λόγο ονομάζει όλη αυτήν την αρχιτεκτονική *«Θεότητα»*, και ενημερώνει ότι αυτή η Θεότητα δεν θα μπορεί να επηρεαστεί από κανέναν και από τίποτα. Ο μόνος τρόπος θα ήταν κάποιος να προσευχηθεί στο όνομά της και να την λατρεύει ως θρησκεία: *«Ξεκινήσαμε μια διαδικασία για να δημιουργήσουμε έναν Θεό. Επομένως, πρέπει να σκεφτούμε και να το κάνουμε με τον καλύτερο τρόπο. Πρόκειται για μια καταπληκτική ευκαιρία»* είπε.⁹³

5.5 Επιπτώσεις στις σχέσεις των δύο φύλων.

Ο τομέας των ρομπότ του σεξ (ή αλλιώς, ρομποτικές κούκλες του σεξ) θεωρείται από πολλούς ότι θα αποτελέσει έναν από τους πιο κερδοφόρους της αγοράς ρομποτικής στο προσεχές μέλλον. Ήδη πολλές κούκλες του σεξ ενσωματώνουν εξελιγμένα τεχνικά χαρακτηριστικά, που σε κάποιες περιπτώσεις αγγίζουν τα όρια της τεχνητής νοημοσύνης.

Εταιρείες όπως η Abyss Creations (sex toys), που ενσωματώνει ηλεκτρονικά χαρακτηριστικά στις συσκευές της, και η True Companion, που αναπτύσσει τη Roxxy (το πρώτο ρομπότ του σεξ), υποστηρίζουν ότι δεν επιδιώκεται η αντικατάσταση της συζύγου ή της φιλενάδας, αλλά η κάλυψη ενός κενού που μπορεί να βιώνουν άνθρωποι που έχουν χάσει πολύ κοντινά τους πρόσωπα.

Σύμφωνα με κάποιες προβλέψεις, σύντομα τέτοια προϊόντα θα εισέλθουν μαζικά στην αγορά, προκαλώντας μαζικές αντιδράσεις πιθανώς, αλλά παράλληλα δημιουργώντας μια τεράστια αγορά για ρομπότ τέτοιου είδους μέχρι το 2050. Πρόσφατο παράδειγμα, είναι του εφευρέτη στο Χονγκ Κονγκ, που δημιούργησε τη δική του γυναίκα- ρομπότ. Το όνομά της είναι Mark 1.

⁹³ *iefimerida.gr, «Μια νέα θρησκεία γεννήθηκε: Ο Θεός της τεχνητής νοημοσύνης», Νοέμβριος 20117*



Εικόνα 12:
Ο Ρίκι Μα μαζί με την σύντροφό του – ανθρωπόμορφο ρομπότ Mark 1.
Πηγή: Reuters

Η Mark 1, μπορεί να απαντά σε προγραμματισμένες ερωτήσεις και ως πρότυπο εμφάνισης έχει μια διάσημη ηθοποιό του Χόλιγουντ, την οποία ο Ρίκι Μα αρνείται να αποκαλύψει (αλλά μάλλον, κρίνοντας από την εμφάνιση, πρόκειται για τη Σκάρλετ Γιόχανσον). Για την κατασκευή της χρειάστηκε 1,5 έτος και το κόστος ήταν 50.000 δολάρια.

Σύμφωνα με την μελέτη του μελλοντολόγου Δρ. Ian Pearson πάνω στον επιστημονικό κλάδο που διατυπώνει προβλέψεις για το μέλλον, υποβοηθούμενος από την επιστήμη και τις τάσεις σε τομείς όπως η βιοτεχνολογία, η ναυτεχνολογία, η τεχνητή νοημοσύνη κλπ, οι ρομποτικοί σεξουαλικοί παρτενέρ το 2025 θα είναι τόσο διαδεδομένοι, όσο είναι και η πορνογραφία σήμερα. Υποστηρίζει ακόμη στην έκθεσή του ότι το "σεξ με ρομπότ" μπορεί να γίνει τόσο δημοφιλές που θα ξεπεράσει την ανθρώπινη επαφή και θα είναι πλήρως αναπτυγμένο το 2050. Πιο συγκεκριμένα, στο δημοσίευσμά του στη Mirror αναφέρεται, πως οι γυναίκες θα προτιμούν περισσότερο την σεξουαλική συνένωση με ρομπότ (ρομποτο-φιλία). Μάλιστα θα είναι τόσο κοινό το σεξ με cyborg - όπως το χαρακτηρίζει, που δεν θα εκπλήσσει. Μπορεί, ισχυρίζεται, στην αρχή να υπάρχει ένας δισταγμός για να κάνουμε σεξ με ρομπότ, σιγά σιγά όμως θα συνηθίσουμε στην ιδέα.

Καθώς μάλιστα στους "μηχανικούς" εραστές θα βελτιώνεται όλο και περισσότερο η τεχνητή νοημοσύνη τους, η αίσθηση και η μηχανική τους συμπεριφορά, ενδεχομένως θα αρχίσουν να αναπτύσσονται ισχυροί συναισθηματικοί δεσμοί που θα εξουδετερώσουν τους ενδοιασμούς και τις επιφυλάξεις, όπως αναφέρει ο Δρ. Pearson. Πολλοί είναι εκείνοι που έχουν εκφράσει ανησυχίες για αυτή την αυξανόμενη τάση για σεξ με ρομπότ, που όπως υπολογίζεται θα αντικαταστήσει τις ανθρώπινες σχέσεις, άρα μπορεί να αντικαταστήσει και την ανθρώπινη αλληλεπίδραση.⁹⁴

Ο Σέρχι Σάντος είναι ο εμπνευστής του «ρομπότ του σεξ» και υποστηρίζει πως έχει σχέδιο για το πως θα αποκτήσει μωρό με το δικό του ρομπότ, τη Σαμάνθα. Ο δημιουργός της υποστηρίζει πως στο μέλλον οι άνδρες δεν θα κρύβουν στις ντουλάπες τους τα ρομπότ τους ή τις κούκλες του σεξ, αλλά θα τις παντρεύονται. «Οι άνθρωποι πριν το καταλάβουν, αυτά τα ρομπότ θα κάνουν τις δουλειές τους, θα παντρευτούν τα παιδιά τους κι έπειτα τα εγγόνια τους», ανέφερε στη «Sun Online».

⁹⁴ Δρ. Θάνος Ασηκτής, Μαρίνα Μόσχα, «Οι γυναίκες θα κάνουν περισσότερο σεξ με ρομπότ παρά με άνδρες το 2025!», Μάιος 2017



Εικόνα 13:

**Ο Sergi Santos μαζί με την σύντροφό του Maritza Kissamitaki κάθονται με το ρομπότ του σεξ, την Σαμάνθα, στο σπίτι τους.
Πηγή: Reuters**

Ο επιστήμονας από την Ισπανία συνέχισε την σκέψη του εξηγώντας πως: «Μπορώ να κάνω αυτά τα ρομπότ να δημιουργήσουν παιδί. Θα αγαπούσα να κάνω παιδί με ένα ρομπότ. Θα μπορούσα να φτιάξω έναν αλγόριθμο με την προσωπικότητα που πιστεύω ότι έχει χρησιμοποιώντας τον εγκέφαλο που ήδη έχω δημιουργήσει και θα έκανα 3D εκτύπωση ένα ρομπότ που θα είναι το παιδί μας. Δεν βλέπω που είναι το κακό». Αυτό φαίνεται να γίνεται αποδεκτό και από την σύζυγό του, που είναι άνθρωπος.⁹⁵

Τα ηθικά ζητήματα που θα προκύψουν από το σεξ μεταξύ ανθρώπων και ρομπότ στο μέλλον, εξερευνά το Ίδρυμα «Υπεύθυνης Ρομποτικής» στη Χάγη, που ξεκινά τον διάλογο για το νομικό πλαίσιο της σεξουαλικής χρήσης μηχανών με «νοημοσύνη». Όμως η εξέλιξη των συστημάτων τεχνητής νοημοσύνης αναμένεται να περιπλέξει τα δεδομένα. Ήδη ψυχολόγοι, φιλόσοφοι και νομικοί, προβληματίζονται για τις συνέπειες που θα έχει στην ιδιωτική αλλά και στη δημόσια σφαίρα η χρήση «υπερβολικά ανθρωπόμορφων» ρομπότ.

Όπως το έθεσε ο CEO της εταιρείας Ματ ΜακΜούλεν, στόχος είναι οι πελάτες να μπορούν στο μέλλον να αναπτύξουν όχι μόνο σαρκική αλλά και συναισθηματική σύνδεση με τα ρομπότ, τα οποία με τον καιρό θα διαμορφώνουν το δικό τους χαρακτήρα. Θα πρέπει να επιτρέπεται το μαστίγωμα ενός ρομπότ με τεχνητή νοημοσύνη και στοιχεία συνείδησης; Θα παραμείνει παράνομο το σεξ με ένα ρομπότ-σκύλο; Για αυτά και για άλλα ακόμα πιο σοβαρά διλήμματα το Ίδρυμα ζητά από τώρα διάλογο για το μέλλον.

Παρόμοια διλήμματα και ανησυχίες είχαν προκύψει και πριν από τρεις δεκαετίες, όταν τα πρώτα ηλεκτρονικά παιχνίδια έκαναν την εμφάνισή τους στις κονσόλες των ανήλικων μαθητών, διχάζοντας την ακαδημαϊκή κοινότητα των κοινωνιολόγων και των ψυχολόγων. Από την μία βρίσκονταν εκείνοι που υποστήριζαν ότι τα βίαια παιχνίδια προκαλούν εξάρτηση και έξαρση της βίας και από την άλλη, όσοι θεωρούσαν ότι η εικονική βία λειτουργεί ως διέξοδος και εκτόνωση, σε ένα ούτως ή άλλως ανύπαρκτο σύμπαν που το e-έγκλημα είναι αφαιρετική μίμηση πράξης, που στερείται εγκληματικής ουσίας.

Κατά καιρούς διάφορες πανεπιστημιακές έρευνες είχαν κάνει την εμφάνισή τους, ενισχύοντας τότε τη μία και τότε την άλλη θεωρία χωρίς τα αποτελέσματα να είναι

⁹⁵ Gazzetta, «Κι όμως ο εμπνευστής του ρομπότ του σεξ, είναι έτοιμος να κάνει παιδί», Μάρτιος 2018

κατηγορηματικά. Όπως όμως εξηγούν οι ειδικοί, το σεξ με αληθοφανή ρομπότ είναι πιο ευαίσθητο θέμα, αφού πρόκειται για *άμεση εμπειρία* που ενδεχομένως να ενισχύει την επιθυμία για παραβατική συμπεριφορά, αντί για την εκτόνωση. Οι ανησυχίες στρέφονται κυρίως στην πιθανότητα οι άνθρωποι με «παράνομες επιθυμίες» να δελεαστούν από τις εμπειρίες τους με τα ρομπότ και να ρισκάρουν να εγκληματίσουν, στο κυνήγι μίας αληθινής εμπειρίας. Στην έκθεση οι ειδικοί αναρωτιούνται για το αν τελικά τα σεξουαλικά ρομπότ θα μειώσουν τα φαινόμενα βιασμού, ξυλοδαρμού και ασέλγειας ή θα τα ενισχύσουν.

Ηλεκτρονικές ψηφοφορίες που διενεργούνται κατά καιρούς, δείχνουν ότι η πλειονότητα των ανδρών και μία μερίδα γυναικών (που δεν ξεπερνά το 33%), εμφανίζεται θετική στη χρήση σεξουαλικών ρομπότ «αλλά με κανόνες». Αρκετοί είναι άλλωστε και εκείνοι που δεν αποκλείουν στο μέλλον άνθρωποι και ρομπότ να ερωτεύονται μεταξύ τους. Τώρα τη σκυτάλη παίρνουν και οι πολιτικοί: τον Φεβρουάριο, επιτροπή του ευρωκοινοβουλίου για τη ρομποτική, συζήτησε το ενδεχόμενο να αναγνωρίσει ταυτότητα στα ρομπότ με τεχνητή νοημοσύνη. Οι πολίτες θα πρέπει να αποφασίσουν τι θα είναι επιτρεπτό και τι απαράδεκτο, υπογραμμίζει ο επικεφαλής της έκθεσης και καθηγητής ρομποτικής στο Πανεπιστήμιο του Σέφιλντ, Νόελ Σάρκι που εκτιμά πως στο μέλλον οι άνθρωποι και τα ρομπότ θα κάνουν «πολύ περισσότερο σεξ» από ότι τώρα.⁹⁶

5.6 Οι συνέπειες της εξέλιξης των ιατρικών υπολογιστικών συστημάτων (θεραπευτική ιατρική χρήση ρομπότ

Σημαντική επίδραση στην καθημερινότητα των ανθρώπων, έχει και η εξέλιξη των ιατρικών ρομποτικών συστημάτων. Στόχος είναι η καλύτερευση των συνθηκών διαβίωσης των ασθενών, αλλά και η βελτίωση των αποτελεσμάτων των επεμβάσεων. Δίνονται επιλεκτικά κάποια παραδείγματα παρακάτω, ώστε να αποδειχτεί η σημαντική επίδραση των ρομποτικών συστημάτων για ιατρική και θεραπευτική χρήση, στην καλύτερευση της ποιότητας ζωής των ανθρώπων, χωρίς να παραλείπονται και οι περιπτώσεις που χρειάζονται περεταίρω έρευνες για την καθιέρωση πιο προηγμένων τεχνικών στον σχεδιασμό και τη λειτουργία των συστημάτων, για την αποφυγή ατυχημάτων και επιπλοκών.

Με το σκεπτικό της καλύτερευσης των συνθηκών διαβίωσης των ασθενών, η εταιρεία τεχνητής νοημοσύνης Google DeepMind, που ανακοίνωσε τη δεύτερη συνεργασία της με το βρετανικό Εθνικό Σύστημα Υγείας (NHS- National Health Service), συνεργαζόμενη με το νοσοκομείο Moorfields Eye στο ανατολικό Λονδίνο, επιδιώκει τη δημιουργία ενός συστήματος machine learning που θα είναι σε θέση να αναγνωρίζει προβλήματα τα οποία μπορεί να απειλήσουν την όραση, απλά και μόνο μέσα από μια ψηφιακή σάρωση του ματιού του ασθενούς. Το project περιλαμβάνει ένα εκατομμύριο ανώνυμες σαρώσεις ματιών, οι οποίες θα χρησιμοποιηθούν για την «εκπαίδευση» ενός αλγορίθμου, ώστε να εντοπίζονται έγκαιρα παθήσεις που μπορούν να απειλήσουν την όραση. Η προηγούμενη, η οποία συνεχίζεται, είναι με το νοσοκομείο Royal Free στο βόρειο Λονδίνο, καθώς και τα Barnet και Chase Farm και έχει να κάνει με τη χρήση εφαρμογής smartphone για την παρακολούθηση της λειτουργίας των νεφρών των ασθενών.⁹⁷

Όσον αφορά τη ρομποτική χειρουργική που θεωρείται ως ένα από τα κορυφαία επιτεύγματα της σύγχρονης τεχνολογίας σε σχέση με την ιατρική, ίσως να μην είναι τόσο ασφαλής όσο πιστευόταν. Νέα μελέτη έδειξε ότι η χρήση τέτοιων μηχανημάτων συνδέεται με αρκετούς θανάτους στις ΗΠΑ. Τα περιστατικά που περιλαμβάνονται στη μελέτη

⁹⁶ Πάνας Σάκκας, «όχι πια σεξ μόνο με ρομπότ, τα μελλοντικά ταμπού του έρωτα με μηχανές», Ιούλιος 2017

⁹⁷ Naftemporiki.gr, «Συνεργασία της Google DeepMind με το βρετανικό Εθνικό σύστημα υγείας», 8 Ιουλίου 2016

(πραγματοποιήθηκε από ερευνητές του University of Illinois at Urbana Champaign, του MIT και του Rush University Medical Center) έχουν να κάνουν με σπασμένα κομμάτια που έπεσαν σε σώματα ασθενών, σπίθες που προκάλεσαν καψίματα και σφάλματα συστήματος που είχαν ως αποτέλεσμα οι εγχειρήσεις να διαρκούν περισσότερο από ό,τι αναμενόταν. Σε κάθε περίπτωση, η αναφορά υπογραμμίζει ότι στατιστικά το ποσοστό είναι μικρό, ωστόσο απαιτούνται επιπρόσθετα και ανανεωμένα μέτρα ασφαλείας. Συνολικά καταγράφηκαν 144 θάνατοι, 1.391 τραυματισμοί και 8.061 δυσλειτουργίες σε ένα σύνολο άνω των 1,7 εκατ. επεμβάσεων μεταξύ του Ιανουαρίου του 2000 και του Δεκεμβρίου του 2013. «Παρά τη διευρυμένη υιοθέτηση ρομποτικών συστημάτων για ελάχιστα επεμβατική χειρουργική, σημειώνεται μη αμελητέος αριθμός τεχνικών δυσκολιών και επιπλοκών κατά τη διάρκεια των επεμβάσεων» αναφέρεται στη μελέτη, όπου υπογραμμίζεται ότι χρειάζεται καθιέρωση πιο προηγμένων τεχνικών στον σχεδιασμό και τη λειτουργία των συστημάτων αυτών για την αποφυγή περαιτέρω συμβάντων στο μέλλον.

Ωστόσο, ο Σάιμον Λέοναρντ, επιστήμονας υπολογιστών του Johns Hopkins University και η ομάδα του, δημοσίευσαν μία έρευνα στο Science Translational Medicine για ένα ρομποτικό βραχίονα που υπόσχεται να ράβει και να συνδέει τμήματα μαλακού ιστού. Αν και ρομποτικοί αυτοματισμοί χρησιμοποιούνται ήδη στη χειρουργική, αξιοποιούνται σε σχέση με στέρεες δομές όπως τα κόκαλα, όπου είναι εύκολο να κρατηθούν σταθερά. Ο μαλακός ιστός κινείται και αλλάζει μορφή όσο γίνονται τα ράμματα, απαιτώντας τις ικανότητες ενός χειρουργού για να γίνουν σωστά. Το ρομπότ STAR (Smart Tissue Automation Robot) απέδειξε τις ικανότητές του σε επεμβάσεις σε γουρούνια. Το όλο σύστημα, που είναι σε θέση να λειτουργεί αυτόνομα, διαθέτει ένα προηγμένο σύστημα 3D imaging και ειδικό υπέρυθρο αισθητήρα για να εντοπίζει ενδείξεις στον ιστό έτσι ώστε να κρατά τη βελόνα στον σωστό δρόμο. Όσον αφορά στην «ποιότητα» των ραμμάτων του, αυτά συγκρίθηκαν με τη δουλειά πέντε χειρουργών που έκαναν την ίδια επέμβαση μέσω τριών διαφορετικών μεθόδων. Αν και χρειάστηκε περισσότερος χρόνος σε κάποιες περιπτώσεις, τα ράμματα κρίθηκαν αντίστοιχα ή καλύτερα αυτών των χειρουργών.

Η εξέλιξη της τεχνολογίας βοηθάει συχνά, στην καλύτερη και σωστότερη διάγνωση, πολλών ασθενειών. Όσο καλύτερα ένας γιατρός μπορεί να κρίνει ότι ένας άνθρωπος πονά και δεν νιώθει καλά, μελετώντας το πρόσωπό του, τόσο καλύτερη και πιο έγκαιρη διάγνωση μπορεί να κάνει. Αλλά, οι γιατροί δεν τα καταφέρνουν πάντα καλά σε αυτό. «Το να ερμηνεύει ένας γιατρός τις εκφράσεις ενός ασθενούς, μπορεί να βοηθήσει στο να διαγνώσει αν έχει πάθει εγκεφαλικό, αν πονάει ή αν εμφανίζει αντίδραση σε κάποιο φάρμακο. Οι γιατροί πρέπει να είναι σε θέση να κάνουν κάτι τέτοιο εξ αρχής», ανέφερε σε σχετική ανακοίνωσή της στη Συνδιάσκεψη για την Αλληλεπίδραση Ανθρώπων-Ρομπότ στη Βιέννη, η Λόρεν Ρικ του Πανεπιστημίου της Καλιφόρνια-Σαν Ντιέγκο, η οποία ανέλαβε με τους συνεργάτες της να δημιουργήσει ένα διαφορετικό ρομπότ. Δημιούργησαν λοιπόν ένα ρομποτικό πρόσωπο, που έχει την ικανότητα να παίρνει εκφράσεις πόνου, φόβου, αηδίας και άλλων συναισθημάτων, προκειμένου να βοηθήσει στην εκπαίδευση των γιατρών στις διαγνώσεις. Το ανθρωποειδές ρομπότ, που φέρει την ονομασία «Φίλιπ Ντικ» (από το όνομα του διάσημου Αμερικανού συγγραφέα επιστημονικής φαντασίας) και δημιουργήθηκε από την εταιρεία Hanson Robotics, σύμφωνα με το «New Scientist», έχει ρεαλιστικό δέρμα στο πρόσωπό του και μπορεί να παίρνει πειστικές εκφράσεις. Εκτός από το ρομπότ, έχει δημιουργηθεί και μια εικονική εκδοχή του, το πρόσωπο ενός ψηφιακού «αβατάρ» σε οθόνη

υπολογιστή, το οποίο προσφέρει μια εναλλακτική μέθοδο εκπαίδευσης στις ιατρικές διαγνώσεις.⁹⁸

Επίσης, ερευνητές από το πανεπιστήμιο του Χάρβαρντ κατασκεύασαν μια ρομποτική μικρογραφία σελαχιού, με σκοπό να μελετήσουν καλύτερα την ανθρώπινη καρδιά, ανοίγοντας ενδεχομένως τον δρόμο ακόμη και για την κατασκευή μίας τεχνητής εκδοχής της. Παράλληλα, όπως αναφέρουν, η χρήση βιολογικού ιστού στο μηχανικό σελάχι θα μπορούσε να βοηθήσει και τον κλάδο των «soft robotics», στόχος του οποίου είναι η ανάπτυξη ρομπότ από εύκαμπτα υλικά, που θα έχουν εξελιγμένες δυνατότητες κίνησης και αποφυγής εμποδίων. Το ρομπότ έχει μικρότερο μέγεθος από ένα νόμισμα των 2 ευρώ, αφού το μήκος του δεν ξεπερνά τα 16,3 χιλιοστά. Επίσης, περιέχει περίπου 200.000 καρδιακά κύτταρα και ζυγίζει περίπου 10,18 χιλιοστά του γραμμαρίου. Τα καρδιακά κύτταρα εξασφαλίζουν την κίνηση των πτερυγίων στο ρομποτικό σελάχι, ώστε να μπορεί αυτό να «κολυμπά» ακολουθώντας μία δέσμη γαλάζιου φωτός. Για να κατασκευάσει το ρομπότ, η ομάδα του Κιτ Πάρκερ χρησιμοποίησε καρδιακά κύτταρα ποντικών, γιατί οι ιστοί τους είναι πιο δυνατοί από αυτούς της ανθρώπινης καρδιάς. Στη συνέχεια, τα τροποποίησαν γενετικά, ώστε να ανταποκρίνονται στο γαλάζιο φως. Το σώμα του μηχανικού σελαχιού κατασκευάστηκε από σιλικόνη, ενώ ο «σκελετός» του από χρυσό. Οι επιστήμονες επέλεξαν το συγκεκριμένο ευγενές μέταλλο επειδή είναι χημικά αδρανές και θα εξασφάλιζε ότι δεν θα καταστρέφονταν τα καρδιακά κύτταρα.⁹⁹

Ακόμα και την υπογεννητικότητα διορθώνει η νανοτεχνολογία. Ένας από τους πλέον κοινούς λόγους σε περιπτώσεις προβλημάτων γονιμότητας είναι σπερματοζώαρια τα οποία δεν μπορούν να κολυμπούν καλά. Οι Μαριάννα Μεντίνα Σάντες, Λούκας Σβαρτς, Όλιβερ Σμιντ και συνάδελφοί τους από το Institute for Integrative Nanosciences (IFW Dresden) στη Γερμανία, αναζήτησαν νέες τεχνικές, δουλεύοντας πάνω σε προηγούμενες έρευνες που είχαν γίνει στον τομέα των μικρο-κινητήρων, ώστε να αναπτύξουν μηχανοκίνητα σπερματοζώαρια – ρομπότ (spermibots) που μπορούν να βοηθήσουν τους όχι ιδιαίτερα ικανούς «κολυμβητές» να φτάσουν στο ωάριο. Η σχετική δουλειά παρουσιάζεται στο Nano Letters. Αυτό που έκαναν ήταν να κατασκευάσουν μικροσκοπικές μεταλλικές προπέλες, τέτοιου μεγέθους, που να επιτρέπεται η εγκατάστασή τους στην ουρά του σπέρματος. Οι κινήσεις τους μπορούν να ελεγχθούν μέσω μαγνητικού πεδίου. Εργαστηριακές δοκιμές έδειξαν ότι τα σπερματοζώαρια μπορούν να κατευθυνθούν έτσι ώστε να οδηγηθούν στον προορισμό τους προκειμένου να γονιμοποιηθεί το ωάριο και εκεί να αφεθούν ελεύθερα.¹⁰⁰

Μια άλλη εξέλιξη στον τομέα της ιατρικής που κάνει πολύ καλύτερη τη ζωή των ανθρώπων επίτευξαν ερευνητές στη Βρετανία, κατασκευάζοντας ένα βιονικό χέρι με τεχνητή νοημοσύνη, το οποίο βλέπει μόνο του και μπορεί έτσι αυτόματα να πιάνει τα αντικείμενα χωρίς σκέψη. Η καινοτόμος προσθετική συσκευή υπόσχεται να δώσει νέες δυνατότητες σε ανθρώπους με ακρωτηριασμένα άνω άκρα. Το τεχνητό χέρι διαθέτει μια κάμερα που βλέπει κάθε αντικείμενο, αξιολογεί το μέγεθος και το σχήμα του, και μετά πράττει την κατάλληλη κίνηση για να το πιάσει. Οι ερευνητές, με επικεφαλής τον δρ Κιανούς Ναζαρπούρ του Τμήματος Βιοϊατρικής Μηχανικής του Πανεπιστημίου του Νιούκαστλ, έκαναν τη σχετική δημοσίευση στο περιοδικό νευρο-μηχανικής, Journal of Neural Engineering «Χρησιμοποιώντας υπολογιστική όραση, δημιουργήσαμε ένα βιονικό χέρι που μπορεί να αντιδράσει αυτόματα. Στην πραγματικότητα, όπως σε ένα πραγματικό χέρι, ο χρήστης μπορεί να

⁹⁸ Onmed.gr, «Ρομπότ βοηθάει τους εκπαιδευόμενους γιατρούς παίρνοντας εκφράσεις πόνου και φόβου», Μάρτιος 2017

⁹⁹ Κώστας Δεληγιάννης, «Ρομποτικό σελάχι με ζωντανά κύτταρα από επιστήμονες του Χάρβαρντ», Ιούλιος 2016

¹⁰⁰ Newsnow.gr, «Μηχανοκίνητα σπερματοζώαρια – ρομπότ για αντιμετώπιση προβλημάτων γονιμότητας», Ιανουάριος 2017

πιάσει ένα φλιτζάνι ή ένα μπισκότο ρίχνοντας απλώς μια γρήγορη ματιά προς την κατεύθυνση του αντικειμένου. Για πρώτη φορά, το βιονικό χέρι αντιδρά χωρίς να μεσολαβήσει κάποια σκέψη», δήλωσε ο Ναζαρπούρ. Εκτός από τη βελτίωση του χεριού, ο απώτερος στόχος των ερευνητών είναι ένα τεχνητό άκρο που θα νιώθει την πίεση και τη θερμοκρασία και το οποίο θα επικοινωνεί απευθείας και αμφίδρομα με τον εγκέφαλο». ¹⁰¹

Τεχνολογίες, όπως για παράδειγμα οι τεχνητές καρδιές και τα βιονικά μέλη, δεν συγκαταλέγονται μόνο στις ιατρικές ανάγκες αλλά και στην ανθρώπινη ενδυνάμωση. Η δημιουργία της τεχνητής καρδιάς έχει ως σκοπό να κρατήσει τον ασθενή ζωντανό, ενώ τα βιονικά μέλη είναι τεχνητά μέλη που κατασκευάστηκαν με σκοπό την αντικατάσταση των μελών του σώματος που χάθηκαν, για οποιοδήποτε λόγο. Πολλοί είναι οι επιστήμονες που ισχυρίζονται ότι βρισκόμαστε ήδη σε μία βιονική εποχή. Το ερώτημα που προκύπτει από όλα αυτά είναι: «τι ποσοστό της ανθρώπινης ύπαρξης μπορούν να αντικαταστήσουν οι μηχανές, ώστε να συνεχίσει ένα τεχνητό πλάσμα να αποκαλείται άνθρωπος;».

Όπως είδαμε από τα παραπάνω, η υπολογιστική νοημοσύνη έχει κάνει άλματα στον τομέα της ιατρικής. Βοηθάει βέβαια, τους ανθρώπους να έχουν μια καλύτερη ζωή, αλλά κάποια από αυτά τα συστήματα έχουν δημιουργήσει ηθικά διλλήματα στην επιστημονική κοινότητα, αλλά και στους μη ειδικούς, απλούς πολίτες. Αυτά τα διλλήματα, που αφορούν όχι μόνο την ιατρική χρήση των υπολογιστικών συστημάτων, αλλά και κάθε άλλη χρήση, θα προσπαθήσω να τα αναπτύξω στο επόμενο κεφάλαιο της εργασίας μου.

5.7 Συνέπειες από την εξέλιξη των στρατιωτικών υπολογιστικών συστημάτων.

Η χρήση μη επανδρωμένων αεροσκαφών από τις ένοπλες δυνάμεις είναι πραγματικότητα και έχει σημειωθεί ήδη σημαντική πρόοδος σε πολλές χώρες. Μία εξ αυτών είναι και το Ισραήλ, όπου η Elbit Systems παρουσίασε πρόσφατα το Seagull: Ένα αρθρωτό, αυτόνομο, μη επανδρωμένο πολεμικό σκάφος (USV- Unmanned Surface Vehicle) πολλαπλών ρόλων. Πρόκειται για ένα σκάφος 12 μέτρων, το οποίο μπορεί να εξοπλιστεί με κομμάτια (modules), τα οποία αντιστοιχούν στη φύση της κάθε αποστολής. Δύο σκάφη μπορούν να ελέγχονται συντονισμένα μέσω ενός συστήματος ελέγχου αποστολής (MCS- Mission Control System) από άλλα, επανδρωμένα πολεμικά πλοία ή από την ακτή. Το σκάφος, όπως υπογραμμίζει η εταιρεία, προσφέρεται ιδιαίτερα για ναρκαλιευτικές επιχειρήσεις, με έμφαση σε δραστηριότητα σε άγνωστες περιοχές. Επίσης, ενδείκνυται για ανθυποβρυχιακές επιχειρήσεις, καθώς αποτελεί μια οικονομική και εύκολα διαθέσιμη επιλογή για αυτή τη δουλειά, χωρίς να κινδυνεύουν επανδρωμένες μονάδες από τα εχθρικά υποβρύχια. Κατά την Elbit, ενισχύει σημαντικά τον διοικητή ενός πολεμικού ή μιας ναυτικής βάσης, παρέχοντάς του γρήγορα αξιόπιστες ανθυποβρυχιακές δυνατότητες για την προστασία περιοχών και στόχων υψηλής αξίας από εχθρικά υποβρύχια ή νάρκες. Όπως αναφέρει η εταιρεία, το σκάφος μπορεί επίσης να χρησιμοποιηθεί και σε αποστολές ηλεκτρονικού πολέμου, καθώς και υδρογραφικής φύσης, με αυτονομία που αγγίζει τις 96 ώρες. ¹⁰²

Οι αμερικανικές ένοπλες δυνάμεις απάντησαν με το μήκους 40 μέτρων «Sea Hunter»: ένα πειραματικό, αυτόνομο, πολεμικό πλοίο, που είναι ειδικά σχεδιασμένο για να κυνηγά εχθρικά υποβρύχια. Το σκάφος αυτό είναι σχεδιασμένο να κινείται αυτόνομα στην επιφάνεια του ωκεανού για διαστήματα 2-3 μηνών τη φορά, χωρίς πλήρωμα ή τηλεχειρισμό. Αυτό το επίπεδο αυτονομίας θα το καθιστούσε εξαιρετικά ικανό κυνηγό υποβρυχίων, με πολύ χαμηλό κόστος σε σχέση με τα σημερινά δεδομένα. Όπως τόνισε ο Ρόμπερτ Γουόρκ, αναπληρωτής

¹⁰¹ ΑΠΕ-ΜΠΕ, «Βιονικό χέρι με τεχνητή νοημοσύνη», Μάιος 2017

¹⁰² Naftemporiki.gr, «Ρομποτικό πολεμικό πλοίο από ισραηλινή εταιρεία», Φεβρουάριος 2016

υπουργός Άμυνας, ελπίζεται πως τέτοιου είδους σκάφη θα αρχίσουν να χρησιμοποιούνται στην δυτικό Ειρηνικό μέσα σε πέντε χρόνια. Οι σχεδιασμοί του Πενταγώνου περιλαμβάνουν μια στρατηγική διευρυμένης χρήσης μη επανδρωμένων σκαφών με όλο και μεγαλύτερη αυτονομία σε έδαφος, θάλασσα και αέρα. Παράλληλα, η Κίνα ενισχύει τον υποβρυχιακό στόλο της, κάτι που εγείρει ανησυχίες στις ΗΠΑ σχε34ετικά με την ασφάλεια των αεροπλανοφόρων και των υποβρυχίων της, που αποτελούν τα κλειδιά της αμερικανικής υπεροχής στον δυτικό Ειρηνικό. Σκοπός είναι η χρήση πλοίων όπως το Sea Hunter σε ένα μεγάλο εύρος αποστολών, πιθανώς και σε ναρκαλιευτικό ρόλο, με ελάχιστη ανθρώπινη παρέμβαση και επίβλεψη.¹⁰³

Στην ανακοίνωση ότι το αμερικανικό πολεμικό ναυτικό σχεδιάζει να αρχίσει να χρησιμοποιεί το drone «Blackwing» στα υποβρύχιά του, επανδρωμένα και μη, προέβη η αμερικανική εταιρεία AeroVironment. Είναι ένα μικρό, εκτοξευόμενο από σωλήνα αεροσκάφος, που μπορεί να εκτοξεύεται κάτω από την επιφάνεια του νερού. Η ανάπτυξη του Blackwing βασίζεται στην επιχειρησιακή εμπειρία της εταιρείας με μικρά μη επανδρωμένα αεροσκάφη (UAS/UAV) και το σύστημα LMAMS (Lethal Miniature Aerial Missile System), Switchblade της εταιρείας. Αυτό σημαίνει πως το drone, πέρα από σκοπούς παρατήρησης και αναγνώρισης, μπορεί να χρησιμοποιηθεί και ως όπλο, καθώς το σκάφος θα φέρει εκρηκτική κεφαλή για να μπορεί να πλήττει αυτό-καταστρεφόμενο, στόχους που μπορεί να προκύψουν. Το Blackwing, σύμφωνα με την εταιρεία, θα μπορεί να χρησιμοποιηθεί για χρήση σε A2/AD (Anti-Access/ Aerial Denial) περιβάλλοντα. Το drone αναπτύχθηκε στο πλαίσιο σχετικού προγράμματος της αμερικανικής Διοίκησης Ειδικών Επιχειρήσεων (AWESUM- Advanced Weapons Enhanced by Submarine UAS against Mobile targets). Διαθέτει εξελιγμένο ηλεκτροοπτικό και υπέρυθρο εξοπλισμό, σύστημα SASSM (Selective Availability Anti-spoofing Module) GPS και ασφαλές data link, προκειμένου να αποτελεί «πολυεργαλείο» για τα αμερικανικά υποβρύχια.¹⁰⁴

Οι πιλότοι μαχητικών αεροσκαφών θεωρούνται από τα κορυφαία στελέχη των ενόπλων δυνάμεων, καθώς πρέπει να συνδυάζουν τεχνική κατάρτιση με υψηλές σωματικές επιδόσεις, όπως όραση, αντοχή κλπ. Οι απαιτήσεις της ιδιότητας του πιλότου μαχητικού είναι πολύ μεγάλες για να καλυφθούν προς το παρόν από μη ανθρώπινα χέρια και για αυτό και προκαλεί αίσθηση το ότι μια τεχνητή νοημοσύνη ήταν σε θέση να νικήσει έναν βετεράνο πιλότο μαχητικών. Ο λόγος για το ALPHA. Αξιοποιώντας τη μεθοδολογία Genetic Fuzzy Tree, Αμερικανοί ερευνητές από την Psibernetix, το Air Force Research Laboratory, το University of Cincinnati, ανέπτυξαν τη συγκεκριμένη τεχνητή νοημοσύνη, για τον έλεγχο μη επανδρωμένων μαχητικών αεροσκαφών (UCAV) σε μάχιμες αποστολές. Ο βετεράνος πιλότος μαχητικών, σμήναρχος εν αποστρατεία Τζιν Λη («Geno») που νικήθηκε σε εικονική αερομαχία από το ALPHA, το περιέγραψε ως την πιο «επιθετική, με γρήγορες αντιδράσεις, δυναμική και αξιόπιστη τεχνητή νοημοσύνη» που είχε δει μέχρι τώρα.

¹⁰³ *Naftemporiki.gr*, «Sea Hunter: Το αμερικάνικο ναυτικό βάφτισε το πειραματικό ρομποτικό ανθυποβρυχιακό του σκάφος», Απρίλιος 2016

¹⁰⁴ *Naftemporiki.gr*, «Μη επανδρωμένα σκάφη για εκτόξευση από υποβρύχια του αμερικάνικου πολεμικού ναυτικού», Μάιος 2016



Εικόνα 14:

Ο σμήναρχος εν αποστρατεία Τζιν Λη στον εξομοιωτή πτήσης ALPHA.

Ο σμήναρχος Λη («Geno») περιέγραψε το ALPHA ως την πιο «επιθετική, με γρήγορες αντιδράσεις, δυναμική και αξιόπιστη τεχνητή νοημοσύνη» που είχε δει μέχρι τώρα.

Πηγή: University of Cincinnati /Lisa Ventre

Το μυστικό της τεχνητής νοημοσύνης είναι το «genetic fuzzy tree», ένα σύστημα λήψης αποφάσεων που μπορεί να προσεγγίζει πολύπλοκα προβλήματα όπως θα το έκανε ένας άνθρωπος, «σπάζοντάς» τα σε μικρότερες υπο-εργασίες, που περιλαμβάνουν τακτικές υψηλού επιπέδου, αποφυγή εχθρού, άνοιγμα πυρών, αμυντικές τακτικές κ.α. Η τεχνητή νοημοσύνη εξετάζει μόνο τις πιο σχετικές μεταβλητές και μπορεί να παίρνει αποφάσεις με πολύ μεγάλη ταχύτητα, υπολογίζοντας τους καλύτερους ελιγμούς μέσα σε «δύσκολα» περιβάλλοντα. Σύμφωνα με δημοσίευμα του Popular Science, το ALPHA είναι 250 φορές ταχύτερο από τον άνθρωπο αντίπαλό του. Δοκιμάστηκε αρχικά απέναντι σε άλλες τεχνητές νοημοσύνες, πριν αντιμετωπίσει τον σμήναρχο Λη, με μακρά εμπειρία τόσο στο κόκπιτ όσο και στη διοίκηση και συντονισμό αερομαχιών. Όταν ο «Geno» ανέλαβε τον έλεγχο αεροσκαφών της «μπλε» πλευράς ενάντια στα «κόκκινα» που ήταν υπό τον έλεγχο του ALPHA, διαπίστωσε ότι υστερούσε σημαντικά, καθώς όχι απλά δεν κατάφερε να προβεί σε κατάρριψη, αλλά έχανε κάθε φορά σε παρατεταμένες εμπλοκές.¹⁰⁵

Ο General Atomics MQ-9 Reaper (Γενικός Ατομικός Αναχαιτιστής MQ-9), μερικές φορές αποκαλούμενος Predator B, είναι ένα μη επανδρωμένο εναέριο όχημα (UAV) ικανό για τηλεχειριζόμενες ή αυτόνομες πτητικές λειτουργίες, που αναπτύχθηκε από την General Atomics Aeronautical Systems (GA-ASI)).¹⁰⁶ Το MQ-9 είναι το πρώτο UAV κυνηγός-δολοφόνος, σχεδιασμένο για επιτήρηση μεγάλης αντοχής, υψηλού υψομέτρου, χάρη στο αναγνωριστικό "Reaper" που του δόθηκε. Το 2006, ο τότε αρχηγός του Γενικού Επιτελείου της Πολεμικής Αεροπορίας των ΗΠΑ, στρατηγός T. Michael Moseley, δήλωσε: *“Έχουμε κινηθεί από τη χρήση UAVs κυρίως στον τομέα της ευφυΐας, της επιτήρησης και αναγνώρισης ρόλων πριν από την επιχείρηση Operation Iraqi Freedom, στην χρήση ενός ρόλου κυνηγού-δολοφόνου με το Reaper.”*

Το MQ-9 είναι ένα μεγαλύτερο, βαρύτερο και πιο ικανό αεροσκάφος από το προηγούμενο Predator MQ-1 της Γενικής Ατμόσφαιρας. Επίσης, μπορεί να ελέγχεται από τα

¹⁰⁵ *Naftemporiki.gr, «τεχνητή νοημοσύνη νίκησε βετεράνο πιλότο μαχητικών σε εικονικές αερομαχίες», Ιούνιος 2016*

¹⁰⁶ *Escutia, Sondra, «4 remotely piloted vehicle squadrons stand up at Holloman», Πολεμική Αεροπορία των ΗΠΑ, Ιανουάριος 2012*

ίδια συστήματα εδάφους που χρησιμοποιούνται για τον έλεγχο των MQ-1. Ο Reaper διαθέτει στροβιλοκινητήρα 950 ίππων, που του επιτρέπει να φέρει 15 φορές περισσότερο ωφέλιμο φορτίο πυροβόλων όπλων και διαθέτει περίπου τρεις φορές την ταχύτητα του MQ-1. Το αεροσκάφος παρακολουθείται και ελέγχεται από αεροσκάφος στο σταθμό ελέγχου εδάφους (GCS). Από τον Μάρτιο του 2011, η Πολεμική Αεροπορία των ΗΠΑ εκπαιδεύει περισσότερους πιλότους για προηγμένα μη επανδρωμένα οχήματα εναέριας κυκλοφορίας από ό, τι για κάθε άλλο σύστημα όπλων. Ο Reaper χρησιμοποιείται επίσης από το Πολεμικό Ναυτικό των ΗΠΑ, τη CIA, την Τελωνειακή και Συνοριακή Προστασία των ΗΠΑ, τη NASA και τους στρατιωτικούς πολλών άλλων χωρών.¹⁰⁷

Ο Αμερικανός γεροστιαστής Λίντσεϊ Γκράχαμ δήλωσε το 2013, ότι τα χτυπήματα με drones έχουν σκοτώσει 4.500 ανθρώπους. Χρησιμοποιώντας την καταμέτρηση του Γραφείου της Διερευνητικής Δημοσιογραφίας, οι ΗΠΑ ξεκίνησαν μεταξύ 416 και 439 χτυπημάτων σε αεροσκάφη στην Υεμένη, το Πακιστάν και τη Σομαλία από τότε που οι ΗΠΑ κατάφεραν με επιτυχία να εξοπλίσουν με επιτυχία ένα Predator MQ-1 πριν από μια δεκαετία. Το 2016, ο πρώην πρόεδρος Μπαράκ Ομπάμα ισχυρίστηκε ότι τα αεροσκάφη σκότωσαν 116 μη εχθρικούς πολεμιστές σε 473 τρομοκρατικές επιθέσεις στο Πακιστάν, την Υεμένη, τη Σομαλία και τη Λιβύη, μεταξύ Ιανουαρίου 2009 και τέλους 2015. Ωστόσο, το Γραφείο Διερευνητικής Δημοσιογραφίας ισχυρίζεται ότι σκοτώθηκαν 380 με 801 άμαχοι κατά το διάστημα αυτό, βάσει ειδησεογραφικών εκθέσεων ερευνητών ΜΚΟ και αφού διέρρευσαν κυβερνητικά έγγραφα και έγγραφα δικαστηρίων. Στο μέλλον, ίσως το νέο MQ-9 να είναι πιο ακριβές στο να σκοτώνει τους επιδιωκόμενους στόχους από τον προκατόχο του.¹⁰⁸

Πολλά ακόμα υπολογιστικά συστήματα έχουν αναπτυχθεί για να μετατρέψουν ένα πιθανό μελλοντικό πόλεμο, σε υπεροχή του κράτους που τα διαθέτει. Όλα τα κράτη, άλλα λιγότερο κι άλλα περισσότερο, έχουν επενδύσει όχι μόνο σε αναγνωριστικά και αμυντικά συστήματα, αλλά και σε συστήματα καταστολής τους εχθρού. Είναι λοιπόν πιθανό να επαληθευτούν οι προφητείες όχι μόνο θρησκευτικών εξεχουσών προσωπικοτήτων, αλλά και επιστημόνων, που διατρανώνουν ότι σε μια πιθανή μελλοντική παγκόσμια συμπλοκή, θα καταστραφεί το μεγαλύτερο μέρος της ανθρωπότητας;

5.8 SUPERINTELLIGENT και ΜΟΝΑΔΙΚΟΤΗΤΑ: Μία ηθική και τεχνολογική πρόκληση.

Αρκετοί συγγραφείς έχουν υποστηρίξει ότι υπάρχει σημαντική πιθανότητα η superintelligence να μπορεί να δημιουργηθεί μέσα σε λίγες δεκαετίες, ίσως ως αποτέλεσμα της αυξανόμενης απόδοσης υλικού και της αυξημένης ικανότητας υλοποίησης αλγορίθμων και αρχιτεκτονικών παρόμοιες με εκείνες που χρησιμοποιούνται από τους ανθρώπινους εγκεφάλους.¹⁰⁹

Φαίνεται πολύ πιο δύσκολο να φτάσουμε σε μία τεχνητή νοημοσύνη ανθρώπινου επιπέδου, απ' ότι σε μία superintelligence. Αν και για να φτάσουμε στην superintelligence μπορεί να διαρκέσει αρκετά χρόνια, το τελικό στάδιο μπορεί να συμβεί γρήγορα. Δηλαδή, η μετάβαση από μια κατάσταση στην οποία έχουμε τεχνητή νοημοσύνη ανθρώπινου επιπέδου, σε μια κατάσταση στην οποία έχουμε μία γεμάτη έννοια της superintelligence, με

¹⁰⁷ Flightglobal.com, «Ο μικρός ταξιδιωτικός χάρτης UAS της USAF απευθύνει έκκληση για τα καμικάζι drones», Μάιος 2016

¹⁰⁸ Terrell Jermaine starr, «The air force is retiring the predator drone for the more deadly reaper», Φεβρουάριος 2017

¹⁰⁹ International Journal of Futures Studies, «How Long Before Superintelligence? », Bostrom 1998, Kurzweil 1999; Moravec 1999

επαναστατικές εφαρμογές, μπορεί να είναι πολύ γρήγορη, ίσως θέμα ημερών και όχι ετών. Αυτή η πιθανότητα μιας ξαφνικής εμφάνισης της superintelligence αναφέρεται ως «singularity hypothesis», δηλαδή *μοναδικότητα*.¹¹⁰

Το αποκορύφωμα της εξέλιξης των υπολογιστικών συστημάτων θα είναι η δημιουργία της Superintelligence, της τεχνητής νοημοσύνης που θα αποτελέσει την υπερδιάνοια του μέλλοντος. Ας δούμε όμως αναλυτικά τι σημαίνουν και πώς ακριβώς αντιμετωπίζουν οι επιστήμονες την superintelligent και τη μοναδικότητα.

5.8.1. Ανάλυση του όρου superintelligence -υπερδιάνοια- και οι προκλήσεις που θα προκύψουν με τον ερχομό τους.

Με τον όρο "superintelligence", ορίζουμε μία "υπερδιάνοια" που ο νους της είναι κατά πολύ ανώτερος από το νου των ανθρώπινων όντων σε όλους τους τομείς. Σύμφωνα με το βιβλίο "Superintelligence", που γράφτηκε από τον Νικ Μπόστρομ (Σουηδός φιλόσοφος στην Οξφόρδη), οι τεχνητές νοημοσύνες θα λειτουργούν και θα εξελίσσονται τελείως διαφορετικά από τα ανθρώπινα όντα. Η ημέρα που οι μηχανές θα καταφέρουν να έχουν ισάξιες διανοητικές δυνατότητες με αυτές των ανθρώπων, πλησιάζει. Σύμφωνα με τον ίδιο, μόλις αυτό επιτευχθεί, η εμφάνιση της υπερδιάνοιας θα γίνει άμεσα, αφού αυτές οι μηχανές θα έχουν την δυνατότητα να εξελίσσονται κατευθείαν και να βελτιώνονται με πολύ γοργούς ρυθμούς.

Εκείνη την ημέρα, όπως ισχυρίζεται ο Μπόστρομ, η πιθανότητα της εξαφάνισης της ανθρωπότητας θα είναι πολύ μεγάλη, και ακόμα και εάν δεν συμβεί αυτό, η ιστορία της θα αλλάξει για πάντα. Στο βιβλίο που αναφέραμε παραπάνω ότι έχει γράψει, αναπτύσσει τους πιθανούς τρόπους εμφάνισης της ΑΙ καθώς και τους πιθανούς τρόπους με τους οποίους τα ανθρώπινα όντα θα κάνουν προσπάθεια να ελέγξουν την υπερδύναμη της.

Χαρακτηριστικά είναι τα δύο παραδείγματα που ανέφερε. Ας υποθέσουμε λοιπόν, ότι οι άνθρωποι αποφάσισαν να κατασκευάσουν την πρώτη ΑΙ, που θα έχει σαν σκοπό να κάνει την ανθρωπότητα ευτυχισμένη. Το πιο πιθανόν είναι ότι αυτή η υπερδιάνοια θα μελετήσει τον ανθρώπινο εγκέφαλο, ώστε να καταλάβει τι συνδυασμό ενδορφινών χρειάζεται για να νιώσει την «ευτυχία». Για να πετύχει τον σκοπό της, θα κλείσει όλοι την ανθρωπότητα σ' ένα δωμάτιο σε ημικωματώδεις κατάσταση και θα τους γεμίζει ενδορφίνες. Στο δεύτερο παράδειγμά του, υποθέτει ότι οι άνθρωποι αποφάσισαν να κατασκευάσουν μια ΑΙ που ο σκοπός της θα είναι πιο απλός. Για παράδειγμα, η αύξηση της παραγωγής συνδετήρων σε ένα εργοστάσιο. Η υπερδιάνοια θα εφεύρει καινοτομίες που αφορούν αυτήν την παραγωγή και στο τέλος θα μετέτρεπε ολόκληρο το εργοστάσιο αλλά και το σύμπαν σε συνδετήρες.

Σύμφωνα με τον ίδιο, η ανθρωπότητα θα πρέπει να βγάλει από το μυαλό της ότι οι υπερδιάνοιες που θα κατασκευαστούν θα είναι τόσο έξυπνες όσο ένας πολύ έξυπνος άνθρωπος. Θα είναι κάτι πολύ ανώτερο από αυτό. Επίσης, παρομοιάζει την διαφορά των μέσων ανθρώπων με τις υπερδιάνοιες με την διαφορά που έχει ο Ουίνστον Τσόρτσιλ (Πρώην Πρωθυπουργός του Ηνωμένου Βασιλείου) με ένα σαλιγκάρι, λέγοντας ότι οι άνθρωποι θα είναι το σαλιγκάρι.

Η λύση που προτείνει ο Μπόστρομ για τον έλεγχο αυτής της τρομερής δύναμης της τεχνητής νοημοσύνης, είναι η εσκεμμένη επιβράδυνση της τεχνολογικής ανάπτυξης, με σκοπό την καλύτερη προετοιμασία της ανθρωπότητας. Για να γίνει σωστά αυτή η προετοιμασία, τα ανθρώπινα όντα θα πρέπει να διαθέσουν πολλή και ώριμη σκέψη καθώς και χρόνο. Η πιθανότητα της έλευσης των υπερ-διανοιών σε 30, 40 ή 50 χρόνια είναι μεγάλη, και

¹¹⁰ Whole Earth Review, Winter issue, «The Coming Technological Singularity. », Vinge 1993, Hanson et al. 1998

το μόνο σίγουρο είναι ότι τότε η ανθρωπότητα δεν θα προλάβει ούτε να αντιληφθεί, ούτε να σκεφτεί, ούτε να αποφασίσει πώς θα πρέπει να αντιμετωπίσει τις τεχνητές νοημοσύνες που θα μας βλέπουν σαν μυρμήγκια.

Είναι πιθανό ότι κάθε τεχνολογία που μπορούμε επί του παρόντος να προβλέψουμε, θα αναπτυχθεί ταχύτατα από την πρώτη άφιξη μιας υπερ-ευφυίας και χωρίς αμφιβολία μαζί με πολλές άλλες τεχνολογίες για τις οποίες είμαστε ακόμα ανίκανοι. Οι προβλεπόμενες τεχνολογίες που μία superintelligence είναι πιθανό να αναπτύξει, περιλαμβάνουν αρχικά μια ώριμη μοριακή κατασκευή. Σ αυτήν, οι τεχνολογικές εφαρμογές είναι πολλές και ποικίλες: πολύ ισχυροί υπολογιστές, προηγμένα όπλα, πιθανώς σε θέση να αποπλίσουν με ασφάλεια μία πυρηνική δύναμη, διαστημικά ταξίδια και ανιχνευτές von Neumann (αυτο-αναπαραγωγή διαστημικών ανιχνευτών), εξάλειψη της γήρανσης και των ασθενειών, λεπτομερή έλεγχο της ανθρώπινης διάθεσης, των συναισθημάτων, των κινήσεων και των κινήτρων, μεταφόρτωση (νευρική ή υπο-νευρωνική σάρωση) ενός συγκεκριμένου εγκεφάλου και εκτέλεση των ίδιων αλγοριθμικών δομών σε έναν υπολογιστή με τέτοιο τρόπο ώστε να διατηρεί την μνήμη και την προσωπικότητα. Επίσης, ανάνηψη των ασθενών με κρουονική¹¹¹ και μια πλήρως ρεαλιστική εικονική πραγματικότητα.¹¹²

Μπορεί να είναι όμως κάθε υπολογιστική διάνοια, μια Superintelligence; Σύμφωνα με τον Μπόστρομ, Superintelligence είναι κάθε διάνοια που έχει κατά πολύ καλύτερες επιδόσεις, ακόμα και από τον καλύτερο ανθρώπινο εγκέφαλο σχεδόν σε κάθε τομέα, συμπεριλαμβανομένης της επιστημονικής δημιουργικότητας, της γενικής σοφίας, και των κοινωνικών δεξιοτήτων. Ο ορισμός αυτός αφήνει ανοικτό το πώς υλοποιείται η superintelligence. Θα μπορούσε να είναι σε ένα ψηφιακό υπολογιστή, σε ένα σύνολο από δικτυωμένους υπολογιστές, καλλιεργημένους φλοιώδεις ιστούς, ή οτιδήποτε άλλο.

Σύμφωνα με αυτόν τον ορισμό, το Deep Blue για παράδειγμα, ένας υπερ-υπολογιστής που αναπτύχθηκε από την IBM και αναμετρήθηκε με τον Κασπάροβ στο σκάκι και τον νίκησε τον Μάιο του 1997, δεν είναι μία superintelligence, δεδομένου ότι είναι έξυπνο μόνο μέσα σε ένα στενό τομέα (το σκάκι), και ακόμα και εκεί δεν είναι πολύ καλύτερο από τους καλύτερους ανθρώπους παίκτες. Ούτε οντότητες όπως εταιρείες ή επιστημονικές κοινότητες μπορούν να θεωρηθούν superintelligences. Αν και μπορούν να εκτελέσουν μια σειρά από πνευματικά κατορθώματα τα οποία κανένας άνθρωπος δεν είναι ικανός να εκτελέσει, δεν έχουν επαρκώς ενοποιηθεί για να θεωρηθούν "διάνοιες" και υπάρχουν πολλοί τομείς στους οποίους αποδίδουν πολύ χειρότερα από ό, τι μονάχα ένας άνθρωπος. Για παράδειγμα, δεν μπορούν να έχουν μια συνομιλία σε πραγματικό χρόνο με την επιστημονική κοινότητα.

Αρκετοί συγγραφείς, επίσης, έχουν υποστηρίξει ότι υπάρχει μεγάλη πιθανότητα η superintelligence να μπορεί να δημιουργηθεί μέσα σε λίγες δεκαετίες, ίσως ως αποτέλεσμα της αυξανόμενης απόδοσης hardware και της αυξημένης ικανότητας υλοποίησης αλγορίθμων και αρχιτεκτονικών παρόμοιων με εκείνες που χρησιμοποιούνται από ανθρώπινους εγκεφάλους. Λαμβάνοντας υπόψη το τεράστιο μέγεθος των συνεπειών της superintelligence, θα ήταν λογικό να λάβουμε σοβαρά υπόψη αυτή την προοπτική, ακόμη και αν κάποιος σκεφτεί ότι υπάρχει μόνο μια μικρή πιθανότητα να συμβεί οποιαδήποτε στιγμή στο κοντινό μέλλον.¹¹³

¹¹¹ Το κίνημα της Κρουονικής, υποστηρίζει την κατάψυξη του νεκρού με την ελπίδα ότι η ιατρική τεχνολογία θα καταφέρει να τον επαναφέρει στη ζωή κάποια στιγμή στο μέλλον.

¹¹² Drexler, Anchor Books: «K. E. Engines of Creation: The Coming Era of Nanotechnology», Νέα Υόρκη 1986

¹¹³ Θεωδωρής Γεωργακόπουλος, «Οι υπερδιάνοιες του μέλλοντος», Σεπτέμβριος 2014

5.8.2. Ανάλυση του όρου *technological singularity* -τεχνολογική μοναδικότητα- και οι φοβείες των ερευνητών.

Τεχνολογική μοναδικότητα (*technological singularity*) ονομάζεται η εμφάνιση μιας υπερευφυΐας με τεχνολογικά μέσα. Αφού οι δυνατότητες μιας τέτοιας ευφυΐας θα ήταν δύσκολο να γίνουν καταληπτές από τον ανθρώπινο νου, η εμφάνιση μιας τεχνολογικής μοναδικότητας αντιμετωπίζεται ως ένας διανοητικός ορίζοντας γεγονότων, πέραν του οποίου τα συμβάντα δεν μπορούν να προβλεφθούν ή να κατανοηθούν. Η συνηθέστερη άποψη των υποστηρικτών της μοναδικότητας είναι ότι, η "ευφυής έκρηξη",¹¹⁴ όπου υπερ-ευφυΐες δημιουργούν συνεχώς γενεές αυξανόμενα ισχυρών νοών, μπορεί να συμβεί ταχύτατα και μπορεί να μην σταματήσει έως ότου οι γνωστικές ικανότητες ενός αντιπροσωπευτικού νου υπερβούν εκείνες ενός οιασδήποτε ανθρώπου.

Ο συγκεκριμένος όρος "μοναδικότητα" ως περιγραφή ενός φαινομένου τεχνολογικής επιτάχυνσης, που ενδεχομένως προκαλεί απρόβλεπτο αποτέλεσμα εντός της κοινωνίας, επινοήθηκε από τον μαθηματικό Τζον φον Νόιμαν, ο οποίος κατά τα μέσα της δεκαετίας του 1950 μίλησε για "*διαρκώς επιταχυνόμενη πρόοδο της τεχνολογίας και αλλαγές στην λειτουργία του ανθρώπινου βίου, που δίνουν την εντύπωση ότι προσεγγίζεται μια ουσιώδης μοναδικότητα στην ιστορία της ανθρώπινης φυλής, πέραν της οποίας οι ανθρώπινες υποθέσεις, όπως τις γνωρίζουμε, δεν θα υφίστανται*". Ο όρος διαδόθηκε από τον συγγραφέα επιστημονικής φαντασίας Vernor Vigne, ο οποίος υποστηρίζει ότι η τεχνητή νοημοσύνη, η ανθρώπινη βιολογική ενδυνάμωση ή τα περιβάλλοντα εγκεφάλου-υπολογιστή μπορούν να είναι πιθανές αιτίες μοναδικότητας. Η έννοια έχει διαδοθεί, επίσης, από μελλοντολόγους όπως ο Ρέι Κέρζουελ, ο οποίος παραθέτει την χρήση του όρου από τον φον Νόιμαν στον πρόλογο του κλασικού βιβλίου του φον Νόιμαν "The Computer and the Brain."

Κάποιοι αναλυτές αναμένουν την εκδήλωση της μοναδικότητας εντός του 21ου αιώνα. Οι εκτιμήσεις βέβαια, ποικίλουν. Ο όρος "τεχνολογική μοναδικότητα" αντανακλά την ιδέα ότι μια τέτοια αλλαγή μπορεί να συμβεί αιφνιδίως και ότι είναι δύσκολο να προβλέψουμε τον τρόπο λειτουργίας ενός τέτοιου κόσμου.¹¹⁵ Δεν είναι σαφές ότι μια έκρηξη ευφυΐας αυτού του είδους θα είναι επωφελής ή όχι, ή ακόμη και υπαρξιακή απειλή, καθώς το ζήτημα δεν έχει εξεταστεί από τους περισσότερους ερευνητές τεχνητής γενικής νοημοσύνης, παρότι το θέμα της φιλικής τεχνητής νοημοσύνης ερευνάται από το Ίδρυμα Μοναδικότητας για την Τεχνητή Νοημοσύνη (το σημερινό Machine Intelligence Research Institute) και το Ίδρυμα "Το Μέλλον της Ανθρωπότητας" (Future of Humanity Institute).

Ο Στίβεν Χόκινγκ, (Βρετανός φυσικός και διευθυντής του ερευνητικού κέντρου του πανεπιστημίου του Κέμπριτζ), ο Φρανκ Βίλτσεκ (φυσικός του MIT) και ο συνάδελφός του Μαξ Τέγκμαρκ, όπως επίσης και ο Στούαρτ Ράσελ (καθηγητής της επιστήμης των υπολογιστών στο πανεπιστήμιο του Μπέρκλι), με ένα άρθρο τους στην εφημερίδα Huffington Post αναλύουν το ενδεχόμενο της ανάπτυξης τεχνητής νοημοσύνης. Σύμφωνα με τους συγγραφείς, η τεχνητή νοημοσύνη έχει ήδη πραγματοποιήσει άλματα, αλλά και επιτεύγματα, όπως για παράδειγμα τα «έξυπνα» αυτοκίνητα που κινούνται χωρίς οδηγό, καθώς και οι προσωπικοί βοηθοί. Όλα αυτά όμως δεν θα είναι τίποτα μπροστά σε αυτά που θα φέρει η μελλοντική τεχνολογική ανάπτυξη της AI.

Σύμφωνα με τους επιστήμονες, δεν υπάρχει καμία φυσική αιτία, που να μπορεί να αποτρέψει μία τεχνητή νοημοσύνη να εκτελέσει υπολογισμούς καλύτερα από έναν ανθρώπινο εγκέφαλο. Αυτό μπορεί να επιτευχθεί αφού τα σωματίδια έχουν την δυνατότητα να

¹¹⁴ David Chalmers, «Singularity, Intelligence Explosion», Singularity Institute for Artificial Intelligence Απρίλιος 2010

¹¹⁵ Yudkowsky, Eliezer (2008), Bostrom, Nick Cirkovic, Milan, «Artificial Intelligence as a Positive and Negative Factor InGlobal Risk», Global Catastrophic Risks (Oxford University Press): 303, ISBN 978-0-19-857050-9

οργανώνονται με αμέτρητους τρόπους, με σκοπό το βέλτιστο αποτέλεσμα. Αμέτρητα είναι τα πιθανά κέρδη από την εξέλιξη της ΑΙ, αφού όλος ο πολιτισμός στηρίζεται στην ιδέα της νοημοσύνης. Με την βελτιστοποίηση αυτής, θα είναι πολύ εύκολο να απαλειφθούν οι ασθένειες, οι πόλεμοι και η φτώχεια. Αυτή η ραγδαία εξέλιξη θα μπορούσε όχι μόνο να μείνει στην ιστορία σαν την μεγαλύτερη ανακάλυψη του ανθρώπου αλλά και να καταστρέψει το μέλλον του.

Κάνοντας μία πρόβλεψη για το μέλλον, τίποτα δεν μπορεί να σταματήσει την διαδικασία της ραγδαίας ανάπτυξης της Τεχνητής Νοημοσύνης, αφού ακόμα και η ίδια θα πάρει την θέση των ανθρωπίνων όντων που την δημιούργησαν. Στη Διεθνή βιβλιογραφία, αυτή η διαδικασία έχει ονομαστεί ως «μοναδικότητα» (singularity). Πολλοί είναι αυτοί που πιστεύουν πως όχι μόνο είναι αναπόφευκτη αλλά και μαθηματικά αποδείξιμη. Το αρχικό ερώτημα που θα θέσουν οι επιστήμονες, είναι το ποιος θα έχει τον έλεγχο αυτής της τεχνητής νοημοσύνης. Αυτό όμως που θα πρέπει να μας απασχολήσει περισσότερο είναι το εάν κάποιος τελικά θα έχει την δυνατότητα να την ελέγχει. Συμπληρώνουν, ότι ακόμα και εάν αυτή η ανάπτυξη μπορεί να έχει τόσες επιδράσεις στην ζωή των ανθρώπων, αυτοί που εξετάζουν το θέμα με την αναγκαία υπευθυνότητα είναι ελάχιστοι.¹¹⁶

¹¹⁶ *Naftemporiki.gr*, «Οι κίνδυνοι της τεχνητής νοημοσύνης», Απρίλιος 2014

ΚΕΦΑΛΑΙΟ 6

ΗΘΙΚΑ ΔΙΛΗΜΜΑΤΑ ΠΟΥ ΠΡΟΚΥΠΤΟΥΝ ΑΠΟ ΤΗΝ ΕΞΕΛΙΞΗ ΤΗΣ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ (Εκτιμήσεις – Συμπεράσματα)

6.1 ΕΙΣΑΓΩΓΗ

Ο Ρενέ Ντεκάρτ, που γεννήθηκε το 1596 στη Γαλλία, έβαλε στη βάση της ανθρώπινης σκέψης το αξίωμα «σκέφτομαι άρα υπάρχω». Τώρα, όμως, στην εποχή της τεχνητής νοημοσύνης, μπορούν οι μηχανές να διεκδικήσουν την αυτονομία της σκέψης τους, με το ίδιο αξίωμα; Οι πρόοδοι στον τομέα της τεχνητής νοημοσύνης είναι ανεξέλεγκτοι. Θα πρέπει άραγε να τεθούν όρια στην έρευνα που θα μπορούσε να οδηγήσει σε απώλεια του ανθρώπινου ελέγχου των συστημάτων που βασίζονται σε υπολογιστές; Καθημερινά, όλο και περισσότερα ερωτήματα ηθικής φύσεως δημιουργούνται από την ραγδαία εξέλιξη των υπολογιστικών συστημάτων. Ερωτήματα που προβληματίζουν όχι μόνο τον απλό κόσμο, αλλά και τους επιστήμονες, που συνεχίζουν να έχουν αντιφατικές απόψεις. Η ταχύτητα της εξέλιξης είναι αποτρεπτική για τη δυνατότητα προ εκτίμησης των αποτελεσμάτων της. Τα διλλήματα αναρίθμητα και οι απαντήσεις ελάχιστες.

Ακόμη κι αν δοθούν αρχικά ορισμένες πρωταρχικές απαντήσεις, και ενώ βαθμιαία ξεκινάει η δημιουργία των αρμόδιων διεθνών φορέων, συνεχίζει να υπάρχει το ερώτημα για το πώς αυτά τα μέτρα που θα συμπεριλαμβάνονται σε αυτές θα γίνουν εφαρμόσιμα. Πολλοί είναι αυτοί που εμπλέκονται στην διαδικασία εφαρμογής των μέτρων, όπως για παράδειγμα η εκάστοτε κυβέρνηση, οι διεθνείς οργανισμοί, οι πολεμικές βιομηχανίες, οι καταξιωμένοι επιστήμονες καθώς και στοχαστές. Τον κύριο λόγο όμως θα έπρεπε να έχουν οι πολίτες, συναποφασίζοντας για το μέλλον τους, ύστερα από επαρκή ενημέρωση.¹¹⁷

Είναι εύλογο ότι στο κοντινό μέλλον η τεχνητή νοημοσύνη θα είναι δεδομένη για περιορισμένες περιπτώσεις, όπως για παράδειγμα να λάβει μία απόφαση ιατρικής περιθάλψης, ή κατανομής πόρων ή να επιλέξει στρατιωτικούς στόχους. Κι αυτό επειδή θα έχει την ικανότητα να συλλέγει γεγονότα και πληροφορίες σε ελάχιστο χρόνο και να παίρνει καλύτερες αποφάσεις απ' ότι ο άνθρωπος, που τις περισσότερες φορές είναι επηρεασμένος από διάφορες καταστάσεις.

Η ανθρώπινη λήψη αποφάσεων επηρεάζεται σε μεγάλο βαθμό από προκαταλήψεις, μεροληψίες, ασυνέπειες, αλλά και από την κόπωση. Βέβαια, κάθε ανθρώπινη μεροληψία, μπορεί να ληφθεί υπόψη και να διορθωθεί μέχρι έναν βαθμό. Επίσης υπάρχουν όρια όσο αφορά το κατά πόσο παράλογη θα είναι μια λανθασμένη απόφαση, ενός ευφυούς και ισορροπημένου προσώπου. Ακόμα και εάν οι ηθικές αποφάσεις του είναι «λιγότερο σωστές», υπάρχει ένα όριο στο πόσο κακές ή λάθος θα είναι αυτές οι αποφάσεις.

Αυτό όμως, δεν συμβαίνει στις περιπτώσεις των μηχανών. Όταν μία μηχανή κάνει κάποιο λάθος, πιθανότατα αυτό το λάθος να είχε πιο απρόβλεπτα αποτελέσματα, απ' ότι θα

¹¹⁷ Νικόλαος Α. Μαυρίδης, «Ρομποτική, Τεχνητή Νοημοσύνη, Αισθητική και Ηθική», τεύχος 69, Οκτώβριος 2015

είχε ένα ανθρώπινο λάθος. Γίνεται επίσης ακόμα πιο δύσκολο να προβλέψουμε το πώς, το γιατί και σε ποιες περιστάσεις η μηχανή θα κάνει κάποιο λάθος και ποιο θα είναι το μέγεθος της αποτυχίας. Οι απλοί αλγόριθμοι θα πρέπει να είναι εξαιρετικά προβλέψιμοι, αλλά ταυτόχρονα να μπορούν να λαμβάνουν αποφάσεις ακόμα και στις πιο παράξενες και ασυνήθιστες συνθήκες. Η μη προβλεψιμότητα των 0.01% αποτυχιών είναι ένας ισχυρός λόγος για να εξεταστεί προσεκτικά από ηθική άποψη, το πώς και το τι μεταφέρουμε στην μηχανή.¹¹⁸

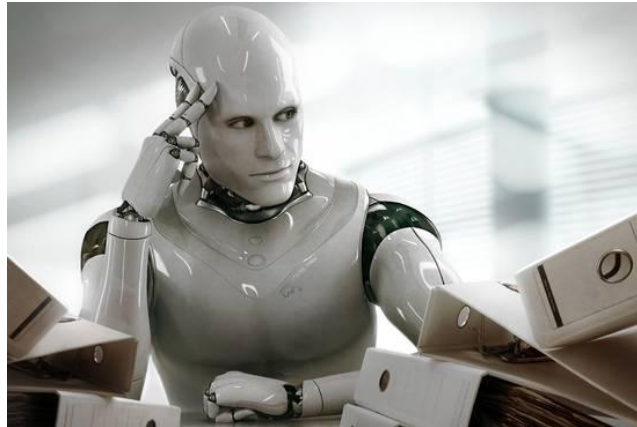
Επίσης, ένα άλλο ερώτημα που δημιουργείται είναι αν τα ρομπότ θα καθιστούσαν τις ζωές των ανθρώπων ευκολότερες, ή αν θα δημιουργούσαν περισσότερα προβλήματα από ότι θα έλυναν. Θα ήταν επικίνδυνα; Θα μπορούσαν να αποτελέσουν απειλή για την ανθρωπότητα; Για να απαντηθούν αυτές οι ερωτήσεις πρέπει να εξεταστούν οι αλληλεπιδράσεις μεταξύ των ίδιων των ευφυών ρομπότ και την επίδραση που μπορεί να έχουν αυτές, στους ανθρώπινους δημιουργούς τους. Παραδείγματος χάριν, εάν επρόκειτο να επιτραπεί σε αισθανόμενες μηχανές να διαπράξουν αδικίες η μία στην άλλη, αυτό θα μπορούσε να προκαλέσει κακό για την ίδια την ανθρωπότητα;

Επιστήμονες, όπως ο Στέφεν Χόκινγκ, ή επιχειρηματίες - επενδυτές, όπως ο Έλον Μασκ, προειδοποιούν ότι οι κίνδυνοι της τεχνητής νοημοσύνης, σίγουρα δεν μπορούν να περάσουν απαρατήρητοι. Συγκεκριμένα ο Στέφεν Χόκινγκ, Καναδοαμερικανός συνιδρυτής της PayPal και ιδρυτής της SpaceX, προειδοποιεί ότι «*ζυπνάμε έναν δαίμονα*», ενώ κι ο Μπιλ Γκέιτς δηλώνει ανήσυχος. Πολλοί επιστήμονες και ηγετικά στελέχη εταιρειών ψηφιακής τεχνολογίας έχουν συνυπογράψει ανοιχτή επιστολή για τον κίνδυνο που ενέχουν οι εξελιγμένες ρομποτικές συσκευές και προειδοποιούν ότι χωρίς έλεγχο το μέλλον του ανθρώπου προδιαγράφεται σκοτεινό. Σαφέστατα η τεχνολογία έχει δώσει λύσεις και έχει βελτιώσει σε αρκετούς τομείς την ζωή του. Τα ρομπότ στην ιατρική, για παράδειγμα, έφεραν μια επανάσταση, αλλά εκεί σαφώς τον έλεγχο τον έχει ο άνθρωπος. Το ερώτημα που τίθεται εδώ είναι, μέχρι ποιου σημείου η υπολογιστική νοημοσύνη θα επιτραπεί να φθάσει και ποιος θα το κρίνει, εντέλει αυτό.

6.2 Ηθική ανάθεση στις μηχανές - Το συναίσθημα ως παράγοντας λήψης ορθής απόφασης.

Όλοι μπορούν να κάνουν ένα είδος ηθικής μεταφοράς σε κάποιο βαθμό, επιτρέποντας στους φίλους, τους συγγραφείς και σε όλα τα δημόσια πρόσωπα να τους επηρεάσουν. Πόσο μάλλον όταν γνωρίζουν ότι αυτοί οι άνθρωποι είναι καλύτερα ενημερωμένοι πάνω σε αυτό το θέμα που τους απασχολεί. Όπως για παράδειγμα κάποιος που έχει καλύτερες δεξιότητες για να διορθώσει μία βλάβη στο αυτοκίνητό τους. Ο καθένας θα μπορούσε να μεταφέρει σκόπιμα ένα ευρύ φάσμα ηθικών ερωτήσεων στους ανθρώπους που ξέρουν ότι μπορούν να απαντήσουν καλύτερα από τους ίδιους ή που είναι απλά καλύτεροι στην παραγωγή ηθικών αξιολογήσεων απ' αυτούς. Φυσικά θα επιλέξουν ανθρώπους που έχουν μία κατά προσέγγιση παρόμοια άποψη των πραγμάτων με αυτούς.

¹¹⁸ Sean O'Heigeartaigh, «*Would you hand over a moral decision to a machine? Why not? Moral outsourcing and Artificial Intelligence*», Πανεπιστήμιο του Όξφορντ, Αύγουστος 2013



Εικόνα 15:

Ο καθορισμός των ηθικών αξιών είναι μια πρόκληση που η ανθρωπότητα έχει αγωνιστεί σε όλη την ιστορία της. Πιστεύετε ότι μπορούμε να σχεδιάσουμε ηθικά ρομπότ, ακόμα κι εάν δεν μπορούμε να συμφωνήσουμε στο τί κάνει έναν άνθρωπο ηθικό;

Όλοι πλέον νιώθουν μια αβεβαιότητα, όσο αφορά τις ηθικές αποφάσεις. Είτε κάποιες από αυτές είναι προφανείς, είτε περίπλοκες. Ποιος έχει τον χρόνο να ερευνησει και να λάβει υγιείς, ηθικές, ενημερωμένες αποφάσεις που αφορούν την διατροφή του, το αντίκτυπο των ενεργειών του στο περιβάλλον, τον φεμινισμό, την πολιτική, το δίκαιο, την κοινωνική ισότητα και άλλα πολλά; Σχεδόν κανείς. Άρα είναι λογικό να αναθέτουν αυτές τις αποφάσεις σε ανθρώπους που ειδικεύονται στο θέμα και μπορούν να λάβουν μία ιδανική για την περίπτωση απόφαση. Το ερώτημα είναι, εάν θα εμπιστεύονταν κάποιος μία "μηχανή" να λάβει μία ηθική απόφαση εξ ονόματός του, όπως θα έκανε με έναν άνθρωπο ειδικό.

Για να απαντηθεί αυτό το ερώτημα, πρέπει πρώτα να εξεταστεί ποιοι παράγοντες προσδιορίζουν την ορθότητα μιας ηθικής απόφασης. Σύμφωνα με μία παραδοσιακή σχολή σκέψης, ο μόνος τρόπος για να ληφθούν σωστές αποφάσεις, είναι η απαλλαγή της λογικής από την επιρροή του συναισθήματος. Τα ερωτήματα όμως είναι εάν τα στοιχεία των νευροεπιστημών συμφωνούν με αυτήν την άποψη, καθώς και εάν ένας ορθός μηχανισμός λήψης αποφάσεων είναι αρκετός για να ληφθούν σωστές αποφάσεις. Κάποιοι υποστηρίζουν ότι χρειάζονται τουλάχιστον ακόμα δύο συστατικά: πρώτον, πληροφορίες και δεύτερον, ένα ξεκάθαρο και πρακτικά υπολογίσιμο ηθικό σύστημα.

Σύμφωνα με τα παραπάνω, το συναίσθημα είναι ενάντια της λογικής; Ο εργάτης Φινέα Γκέιτζ, το 1848 τραυματίστηκε λόγω μίας μεγάλης έκρηξης, καθώς μια μεγάλη μεταλλική ράβδος εισχώρησε στο κρανίο του. Το όνομα του Γκέιτζ έμεινε ιστορικό στις νευροεπιστήμες. Κλινικά αξιοπερίεργο είναι ότι ο Γκέιτζ γλίτωσε από το ατύχημα, διατηρώντας μάλιστα άθικτες τη γενική νοημοσύνη του, την αντίληψη, τη μακροχρόνια μνήμη και έδειξε αλλαγές μόνο σε ότι αφορά την προσωπικότητά του.

Πολλά ανθρώπινα όντα με παρόμοιες εγκεφαλικές βλάβες, αποτελούν σπουδαία στοιχεία για την έρευνα του ανθρώπινου εγκεφάλου και της σύνδεσής του με την νόηση και τον τρόπο αντίδρασης. Αξιοσημείωτα είναι τα αποτελέσματα των ερευνών από τα οποία προκύπτει ότι οι άνθρωποι με βλάβες σε σημεία του εγκεφάλου που επηρεάζουν το συναίσθημα και όχι την λογική, στο τέλος αποκτούν σοβαρό πρόβλημα ακόμα και σε λογικές λειτουργίες, όπως για παράδειγμα η λύση μαθηματικών προβλημάτων. Κατά τη διάρκεια επίλυσης προβλημάτων, γίνονται και βήματα που δεν έχουν να κάνουν με τη λογική, αλλά με την παραγωγική διαδικασία της νόησης που κατευθύνει στη επίλυση, σαν η επίλυση απλά να "μας έρχεται" στο μυαλό. Σε αυτό το σημείο μοιάζει να επεμβαίνουν οι λειτουργίες από περιοχές του εγκεφάλου, που επηρεάζουν το συναίσθημα. Από αυτό προκύπτει ότι η

"λογική νοητική διαδικασία", συμπληρώνεται από ορισμένες συναισθηματικές διαδικασίες. Εν κατακλείδι, το συναίσθημα τις περισσότερες φορές όχι μόνο δεν εμποδίζει την λογική λειτουργία, αλλά είναι απαραίτητο για την λήψη ορθότερων αποφάσεων.

Φυσικά, παρ' όλα αυτά υπάρχουν «συναισθηματικές και άλλες συστηματικές στρεβλώσεις (cognitive biases)» που κάνουν τους ανθρώπους να επιλέγουν την σωστή απόφαση, ακόμα κι αν μπορούν να δικαιολογηθούν σε κάποιες περιπτώσεις. Τα πράγματα λοιπόν, σε σχέση με την εμπλοκή συναισθηματικών διαδικασιών στην λήψη ορθών αποφάσεων, δεν είναι τόσο απλά. Εξαρτώνται πολύ από το πώς κανείς οριοθετεί την λογική και το συναίσθημα, που κατά άλλους δεν αποτελούν διακριτές κατηγορίες. Άρα, από τα παραπάνω φαίνεται ότι δεν υπερτερούν τα συστήματα υπολογιστικής νοημοσύνης σε σχέση με τον άνθρωπο όσον αφορά την επιλογή της σωστής απόφασης. Το σίγουρο είναι ότι αυτό δεν συμβαίνει λόγω της «έλλειψης συναισθηματικών στρεβλώσεων».¹¹⁹

Ο Antonio Damasio (Διευθυντής στο Τμήμα Νευρολογίας στο ιατρικό κολέγιο του Πανεπιστημίου της Ιονα και καθηγητής στο Ινστιτούτο Βιολογικών Ερευνών Salk La Jolla του Σαν Ντιέγκο), στο βιβλίο του «Το λάθος του Καρτέσιου», γράφει χαρακτηριστικά: «*Το να έχει ένας οργανισμός νόηση σημαίνει ότι σχηματίζει νευρωνικές αναπαραστάσεις που μπορούν να καταστούν εικόνες, να επεξεργαστούν με μια διαδικασία που αποκαλούμε σκέψη και τελικά να επηρεάσουν τη συμπεριφορά βοηθώντας στην πρόβλεψη του μέλλοντος, το σχεδιασμό και την επιλογή της επόμενης δράσης*»¹²⁰.

Επίσης στο βιβλίο του αυτό, αποδεικνύει με πειράματα και επιχειρήματα, ότι η λογική δεν μπορεί να λειτουργήσει σωστά αν δεν λειτουργεί φυσιολογικά και το συναίσθημα. Με άλλα λόγια, υποστηρίζει ότι η λογικότητα έχει μια συναισθηματική βάση και ότι λογικότητα και συναίσθημα διασυνδέονται με τη λειτουργία του σώματος που παράγει τις συγκινήσεις. Διατυπώνει μάλιστα, την πλήρη αντίθεσή του στον Καρτέσιο, που φαντάστηκε τη νόηση εντελώς αποκομμένη από το σώμα. Βεβαιώνει ότι ύστερα από ψυχολογικές δοκιμασίες, κατέληξε στο συμπέρασμα, ότι ένας εγκέφαλος διατηρούμενος ζωντανός, αλλά αποκομμένος από το υπόλοιπο σώμα του, δε θα μπορούσε να λειτουργήσει όπως ένας φυσιολογικός εγκέφαλος με πλήρες σώμα. Δε θα είχε φυσιολογική νόηση. Γι' αυτό, ανασκεύασε το περίφημο "σκέφτομαι άρα υπάρχω" του Καρτέσιου, σε "αισθάνομαι άρα υπάρχω". "Ο νους πρέπει όχι μόνο να μετακινηθεί στο βασίλειο του βιολογικού ιστού, αλλά επίσης και να σχετιστεί με έναν συνολικό οργανισμό, ο οποίος κατέχει ένα ολοκληρωμένο σώμα και εγκέφαλο, που αλληλεπιδρούν πλήρως με το φυσικό και κοινωνικό περιβάλλον".¹²¹ Σύμφωνα με αυτή την άποψη, συμπεραίνει κανείς ότι για να απαντηθεί το ερώτημα αν ένα ρομπότ μπορεί να πάρει λογικές αποφάσεις, πρέπει πρώτα να απαντηθεί το ερώτημα αν ένα ρομπότ μπορεί να έχει συναίσθημα.

Εδώ λοιπόν, γεννιέται το ερώτημα αν θα μπορέσουν ποτέ οι μηχανές να "έχουν συναισθήματα", στις περιπτώσεις που απαιτείται και τεχνητή και συναισθηματική νοημοσύνη από αυτές. Ισχύει τελικά ότι το συναίσθημα αποτελεί μόνο ανθρώπινο χαρακτηριστικό; Από ότι φαίνεται, τόσο οι μηχανές της περιόδου που διανύουμε, όσο και αυτές που πρόκειται να εμφανιστούν στο μέλλον, έχουν κάποιες ικανότητες «συναισθηματικής νοημοσύνης». Για παράδειγμα, υπάρχουν αλγόριθμοι που διαβάζουν τις αντιδράσεις του προσώπου των ανθρώπων ή ακόμα και την φωνή τους και τους κατηγοριοποιούν ανάλογα (λυπημένος, αγχωμένος κλπ.). Υπάρχουν και ψεύτικες προσωπικότητες ή υπολογιστικά συστήματα που γελάνε ή τρομάζουν. Χαρακτηριστικό παράδειγμα είναι το ιαπωνικό παιχνίδι τσέπης που

¹¹⁹ Νικόλαος Μαυρίδης, «Τι χρειάζεται για να δοθεί σε κάποιον η άδεια φόνου;», Μάρτιος 2014

¹²⁰ Antonio Damasi, «Το λάθος του Καρτέσιου. Συγκίνηση, Λογική και Ανθρώπινος Εγκέφαλος.» σελ. 144, Μετάφραση Κώστας Παπακωνσταντίνου, Δεκέμβριος 2000

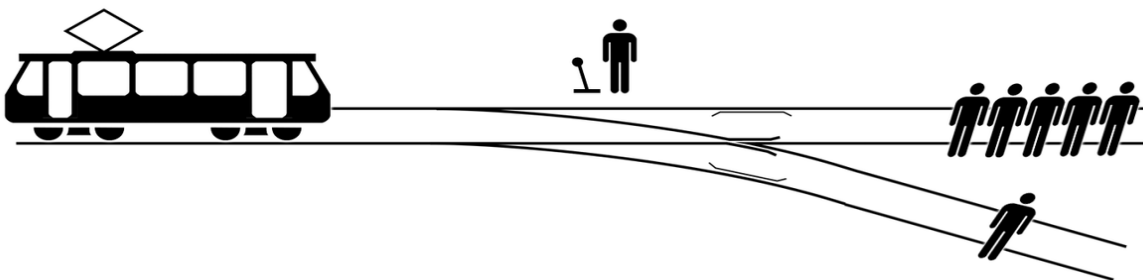
¹²¹ ΝΙΚΟΣ ΒΑΪΔΑΚΗΣ (Αναπλ. καθ. Ψυχιατρικής Κλινικής στο Παν/μιο Αθηνών), «Το λάθος του Καρτέσιου», σελ. 362, ΒΙΒΛΙΟΘΗΚΗ – Μάρτιος 2002

ονομάζεται Tamagochi. Έχει την μορφή ανθρώπου και για να καταφέρει να αναπτυχθεί, πρέπει να είναι χαρούμενο και να το φροντίζουν. Τις περισσότερες φορές οι αλγόριθμοι μπορούν να αναγνωρίσουν ακόμα και ελάχιστες ενδείξεις συναισθημάτων πολύ πιο εύκολα από ότι τις αναγνωρίζουν οι άνθρωποι.

Εδώ όμως, γεννιέται ένα άλλο ερώτημα. Όταν ένα υπολογιστικό σύστημα έχει την ικανότητα αναγνώρισης συναισθημάτων ή έστω παρουσιάζει ότι έχει συναίσθημα, πώς μπορούμε να γνωρίζουμε εάν όντως έχει αληθινά συναισθήματα; Σύμφωνα με τον Νικόλαο Μαυρίδη, η απάντηση του ερωτήματος έχει αντικρουόμενες απόψεις. Κατά αυτόν όμως δεν χωράει αμφισβήτηση. Για παράδειγμα, τα συναισθήματα των άλλων ανθρώπων ή ακόμα και του σκύλου μας, τα συμπεραίνουμε από τις εκφράσεις τους και την συμπεριφορά τους, χωρίς να έχουμε μια άμεση, πρωτογενή πρόσβαση στην εσωτερική κατάσταση οποιουδήποτε άλλου ανθρώπου πλην του εαυτού μας. Γεννιέται λοιπόν το ερώτημα, γιατί να υπάρχει διαφορά στο χαμόγελο ενός ρομπότ; Καταλήγοντας, φτάνει στο συμπέρασμα ότι αν δεχτούμε ότι ο συνάνθρωπος μας, στου οποίου τα συναισθήματα δεν έχουμε πρωτογενή πρόσβαση, αλλά μόνο δευτερογενή, έχει συναισθήματα, τότε έπεται ότι και οι μηχανές μπορούν να έχουν όχι μόνο συναισθηματική νοημοσύνη, αλλά και συναισθήματα, σύμφωνα με την παραπάνω θεώρηση.

Τα πρωταρχικής σημασίας ερωτήματα για την συλλογιστική ηθική είναι δύσκολα, και διχάζουν την κοινή γνώμη. Τι θα συνέβαινε σε περίπτωση που μια πτήση, χωρίς κυβερνήτη, με έναν ανειδίκευτο, σχεδόν αναίσθητο πιλότο, ζητήσει έγκριση προσγείωσης σε ένα πολυπληθές αεροδρόμιο; Το υπολογιστικό σύστημα θα την δεχτεί ή προσπαθώντας να αποφύγει ένα μεγαλύτερο ατύχημα με περισσότερους νεκρούς, τελικά θα αποφασίσει να καταστρέψει το αεροπλάνο; Σχεδόν ίδιο με αυτό, είναι το κλασικότατο «πρόβλημα του τρόλεϊ» της Φιλίππα Φουτ που αναλύεται παρακάτω.

Το παραπάνω νοητικό πείραμα δεοντολογίας, τέθηκε για πρώτη φορά από την Philippa Foot και αναλύθηκε διεξοδικά από τους Judith Jarvis Thomson, Peter Unger και Frances Kamm. Εκτός του τομέα της παραδοσιακής φιλοσοφικής συζήτησης, το πρόβλημα του τρόλεϊ, έχει να κάνει και με τους τομείς της γνωστικής επιστήμης και της νευρο-ηθικής. Το πρόβλημα έχει ως εξής: Μία άμαξα που βρίσκεται εκτός ελέγχου, κατευθύνεται προς μια ομάδα τεσσάρων εργατών. Σε μια παρακείμενη γραμμή, υπάρχει ένας ακόμα εργαζόμενος. Ο θεατής έχει την δυνατότητα να γυρίσει έναν μοχλό και να αλλάξει την πορεία της άμαξας. Θα επέμβει κι αν ναι, θα στείλει την άμαξα στους πέντε εργάτες, που είναι και η αρχική πορεία της άμαξας, ή στον ένα ώστε να σκοτωθούν λιγότεροι άνθρωποι.¹²²



Εικόνα 16:

Το πρόβλημα του τρόλεϊ. Ποια είναι η πιο ηθική επιλογή; Ποια θα ήταν άραγε η επιλογή ενός ανθρώπου και ποια ενός ρομπότ;

¹²² Philippa Foot, «The Problem of Abortion and the Doctrine of the Double Effect in Virtues and Vices», Oxford: Basil Blackwell, 1978

Συνεχίζοντας με το θέμα της λήψης ορθών αποφάσεων, είναι αναγκαίο να εξεταστεί η ύπαρξη επαρκών πληροφοριών, σχετικών με την απόφαση που πρέπει να ληφθεί. Η τροφοδοσία σωστών δεδομένων και ο «νοητικός μηχανισμός», είναι τα απαραίτητα χαρακτηριστικά για την λήψη της σωστής απόφασης. Επίσης, το πιθανότερο είναι ότι το μεγαλύτερο κομμάτι αυτής της πληροφορίας, βρίσκεται αποκλειστικά στις εξωτερικές πληροφορίες, όπως για παράδειγμα το διαδίκτυο και όχι μόνο στις αισθήσεις ή τη μνήμη. Στην ταινία του Ρόμποκοπ, οι μηχανές μπορούν να φτάσουν άμεσα στην πληροφορία, σε αντίθεση με τον άνθρωπο που χρειάζεται «ειδικό ενδιάμεσο – Human machine interface». Για αυτό τον λόγο οι μηχανές υπερτερούν των ανθρωπίνων όντων. Επειδή όμως οι πληροφορίες σήμερα είναι σε μορφή που δεν μπορεί να κατανοηθεί από τις μηχανές, οι άνθρωποι θεωρούνται το ίδιο ικανοί με τις μηχανές και το αντίστροφο.

Τέλος, πρέπει να εξεταστεί η δυνατότητα ύπαρξης ενός ξεκάθαρα και υπολογίσιμου ηθικού συστήματος, για τη λήψη ορθής απόφασης. Τα ερωτήματα που προκύπτουν είναι αρκετά. Τι θεωρείται σωστό και τι λάθος, τι κακό και τι καλό για την λήψη της σωστής απόφασης; Είναι αποδεκτό σαν ύπαρξη το "απόλυτο κακό" ή "απόλυτο καλό", ή το "καλύτερο" ή το "χειρότερο"; Εάν είναι, τότε με ποιον τρόπο θα τοποθετήσουμε τα όρια της έρευνας για τις επιπτώσεις που μπορεί να έχει μία πράξη; Επίσης ένα άλλο ερώτημα που θα πρέπει να απαντηθεί είναι ως προς ποιον θα κρίνουμε αν είναι καλό ή κακό; Ως προς τον εαυτό μας, το παιδί μας, την χώρα μας ή γενικά για την ανθρωπότητα;

Σε αυτά που ένας άνθρωπος δεν θα ήθελε να του συμβούν, υπάρχει μία κοινή πανανθρώπινη συμφωνία, και για αυτόν τον λόγο θεσπίστηκαν τα ανθρώπινα δικαιώματα. Τι θεωρεί όμως θετικό για την ζωή του ένας άνθρωπος; Να έχει γνώσεις, χρήματα, δύναμη, αγάπη; Μπορεί, για παράδειγμα να ισχυριστεί κανείς ότι, ένα υπολογιστικό σύστημα δεν διαφοροποιεί την οπτική γωνία του "Κοινού Καλού", απλά προσπαθεί να αποφύγει τα κοινώς αποδεκτά "κακά". Για παράδειγμα εάν ένας κακοποιός δολοφονηθεί από έναν πολίτη, πώς θα υπολογιστούν οι συνέπειες των πράξεών του και ποιο σύστημα «ηθικής συλλογιστικής (ethical reasoning)» θα χρησιμοποιηθεί από το ρομπότ;

Εκτός όλων αυτών, υπάρχει και κάτι άλλο που με αυτό πρέπει να εφοδιαστεί ένα υπολογιστικό σύστημα. Αυτό της «πρακτικής υλοποιησιμότητας». Πολλά ρομπότ στηρίζουν τα ηθικά τους συστήματα στην δυνατότητα της πρόβλεψης. Με άλλα λόγια, θα πρέπει να επιλέγεται εκείνη η ενέργεια που έμμεσα ή άμεσα θα αποφέρει, σύμφωνα με τις πιθανότητες το πιο ορθό, ηθικά, αποτέλεσμα. Είναι αναγκαίο λοιπόν, να εξετάζεται, πώς μία πράξη θα επηρεάσει μελλοντικά τις πράξεις άλλων ανθρώπινων όντων. Αν και φαντάζει ιδανικό, τις περισσότερες φορές είναι πολύ δύσκολη η πρόβλεψη των δευτερογενών συνεπειών από ένα υπολογιστικό σύστημα. Κύριοι λόγοι είναι οι ανεπαρκείς πληροφορίες και η υπολογιστική πολυπλοκότητα. Για αυτόν τον λόγο τα πιο απλά συστήματα, τα οποία δέχονται απλές εντολές κάτω από προβλέψιμες συνθήκες, έχουν καλύτερη απόδοση.¹²³

6.3 Ηθικά διλήμματα που προκύπτουν από την χρήση υπολογιστικών συστημάτων.

Τα υπολογιστικά συστήματα, ως εργαλεία, τονίζουν μια από τις πιο ισχυρές πτυχές της ανθρωπότητας: την ικανότητά της να ενεργεί αποτελεσματικά στον κόσμο και να τον διαμορφώνει σύμφωνα με τις επιθυμίες της. Το προνόμιο αυτό της ανθρωπότητας παραμένει, όσο διασφαλίζεται ότι δεν θα υποβαθμιστούν, ξεπεραστούν και διαγραφούν οι επιθυμίες των άλλων. Οι τεχνολογίες είναι επεκτάσεις του ατόμου, κωδικοποιημένες σε μηχανές και

¹²³ Νικόλαος Μαυρίδης, «Τι χρειάζεται για να δοθεί σε κάποιον η άδεια φόνου;», Μάτιος 2014

υποδομές, στα πλαίσια της γνώσης και της δράσης. Οι υπολογιστές δεν είναι εδώ για να δώσουν όλες τις απαντήσεις, αλλά για να επιτρέψουν να τεθούν νέα ερωτήματα, με νέους τρόπους, στο σύμπαν.

Οι προβληματισμοί είναι σύνθετοι και έχουν πολλές προεκτάσεις. Συνοψίζονται όμως στο ερώτημα αν υπάρχει ηθική γενικά, στην Υπολογιστική Νοημοσύνη. Ένα σημαντικό ερώτημα που προκύπτει επίσης είναι, εάν οφείλει το ανθρώπινο γένος, να συμπεριφέρεται στις μηχανές σαν να είναι ένα ανώτερο είδος, ή σαν ένα απλό εργαλείο. Όπως είναι γενικά αποδεκτό, οι μηχανές, τα ρομπότ και η Τεχνητή Νοημοσύνη, δημιουργήθηκαν από τον άνθρωπο προκειμένου να βοηθήσουν την ανθρωπότητα να προοδεύσει. Όπως υπάρχει χρήση για τα ρομπότ στο να συναρμολογούν αυτοκίνητα, να κουβαλάνε κιβώτια και να αποπλίζουν βόμβες, άλλο τόσο μπορούν να βοηθήσουν σε πολλούς τομείς της επιστήμης. Υπάρχουν όμως δύο πολύ σημαντικά και καθοριστικά χαρακτηριστικά των ανθρώπων που δεν τα έχουν τα υπολογιστικά συστήματα, που τα κάνει να υστερούν σε σύγκριση με τον άνθρωπο, κι αυτά είναι ο φόβος και ο πόνος. Αυτός είναι ο λόγος που οι επιστήμονες προσπαθούν να τα μεταδώσουν με διάφορους τρόπους στα ρομπότ, δημιουργώντας όμως, ηθικά διλλήματα.

Ο φόβος είναι αυτός που προστατεύει τους ανθρώπους από το να κάνουν λανθασμένες κινήσεις και ο πόνος είναι αυτός που όταν κάνουν την λάθος κίνηση, τους μαθαίνει να προστατεύουν τον εαυτό τους και να μην επαναλάβουν το λάθος αυτό. Αν ένα ρομπότ βιώνει φόβο και πόνο, τότε οι αποφάσεις του θα επηρεάζονται από αυτούς τους δυο παράγοντες. Θα μπορεί να αποφύγει επικίνδυνες καταστάσεις αλλά και να αντιδρά αυτόματα σε άλλες σαν να έχει δικά του αντανακλαστικά. Επίσης, μελετώντας οι επιστήμονες τον πόνο και τον φόβο στα ρομπότ, μπορούν να κατανοήσουν σε μεγαλύτερο βαθμό τον ανθρώπινο εγκέφαλο, χωρίς μάλιστα να χρειαστεί να πειραματιστούν σε ανθρώπους ή άλλους ζωντανούς οργανισμούς.¹²⁴

Κάποιοι όμως ισχυρίζονται, ότι αυτά τα πειράματα δεν πρέπει να γίνονται γιατί τα έξυπνα υπολογιστικά συστήματα είναι σαν ζωντανοί οργανισμοί. Για τους λόγους που έχουν αναλυθεί σε προηγούμενα κεφάλαια, μερίδα επιστημόνων αντιμετωπίζει τα ρομπότ όπως μια ζωική ή ανθρώπινη οντότητα. Υπάρχει βέβαια και η αντίθετη άποψη που λέει ότι ένα ρομπότ είναι απλά, ένα τεχνολογικό επίτευγμα. Είτε λοιπόν το σπρώχνει κάποιος άνθρωπος, είτε του βάζει εμπόδια στον δρόμο του, όλα αυτά γίνονται για την βελτίωση της τεχνολογίας και δεν υπάρχει τίποτα το ανήθικο σε αυτό.

Ένα γνωστό παράδειγμα ο "Atlas" της Boston Dynamics (ανήκει εδώ και καιρό στη Google), που είναι ένα από τα πλέον πολλά υποσχόμενα μοντέλα ανδροειδών που υπάρχουν αυτή τη στιγμή. Το νέο βελτιωμένο μοντέλο έχει πολύ ανώτερες δυνατότητες αυτόνομης κίνησης σε σχέση με το προηγούμενο, αφού φέρει την πηγή ενέργειας στο εσωτερικό του, χωρίς να χρειάζεται να είναι συνδεδεμένο με κάποια εξωτερική πηγή. Μπορεί να σηκώνει και να μετακινεί κιβώτια, ενώ είναι ικανό να τα βγάζει πέρα με πιθανή παρενόχληση από ανθρώπους, παραμένοντας όρθιο ακόμα και μετά από σπρωξίματα, χτυπήματα με μπαστούνι του χόκεϊ κλπ. Μάλιστα, είναι σε θέση να ξανασηκώνεται ακόμα και όταν, μετά από δυνατό σπρωξίμο από πίσω, πέφτει στο πάτωμα προς τα εμπρός. Είναι σχεδιασμένο να κινείται τόσο σε εσωτερικούς, όσο και σε εξωτερικούς χώρους, ακόμα και σε δύσκολα εδάφη, όπως χιόνια, και να προβαίνει μόνο του σε διορθώσεις για να συνεχίσει να κινείται. Σημειώνεται πως ο Atlas, ύψους 1,75 μέτρων και 82 κιλών, είναι ηλεκτροκίνητος και κινεί τα μέλη του με υδραυλικά συστήματα. Όσον αφορά στην αντίληψή του, προκύπτει χάρη σε αισθητήρες στο σώμα και τα πόδια, LIDAR, αισθητήρες στο κεφάλι για αποφυγή εμποδίων, αξιολόγηση του εδάφους, πλοήγηση κ.α.¹²⁵

¹²⁴ Χρήστος Ιβάνοβιτς, «Ηθική στην τεχνητή νοημοσύνη», Αναδημοσίευση από Γιάννη Σπυριδάκη, Ιούνιος 2016

¹²⁵ *Naftemporiki.gr*, «Ένα ρομπότ που στέκεται στα πόδια του», Φεβρουάριος 2016

Σύμφωνα με μια μερίδα επιστημόνων, οι δοκιμές ισορροπίας και ευελιξίας στην περίπτωση του Άτλα, έπρεπε να γίνουν με συγκεκριμένο τρόπο. Το ότι η μηχανή μοιάζει με άνθρωπο, δεν σημαίνει ότι είναι και στην πραγματικότητα. Η ανθρώπινη μορφή του, δεν μειώνει την επίδοσή του, ούτε την κάνει σαδιστική ή μαζοχιστική. Είναι αποδεκτές εξάλλου, οι δοκιμές σύγκρουσης αυτοκινήτων με αερόσακο όπου χρησιμοποιούνται οι καλωδιωμένες κούκλες.



Εικόνα 17:

Το ανθρωποειδές ρομπότ Atlas, της Boston Dynamics, ένα από τα πιο εξελιγμένα ρομπότ που κινείται σαν άνθρωπος.

Υπάρχει η άποψη λοιπόν ότι όσο προηγμένη τεχνολογία κι αν έχουν τα υπολογιστικά συστήματα, αυτό δεν είναι λόγος να τους συμπεριφερόμαστε σαν ανώτερα είδη. Η Τεχνητή Νοημοσύνη είναι μια ακόμα τεχνολογία και ασχέτως αν κάποιες εφαρμογές της έχουν ανθρώπινη μορφή, δεν παύει να είναι μια τεχνολογία που βασίζεται σε κυκλώματα και μικροτσίπ και όχι σε ζωντανά κύτταρα και οργανισμούς.

Η αντίθετη αντιμετώπιση από μέρους των ανθρώπων, που προέρχεται από την ανθρώπινη τάση που θεωρεί τις μηχανές σαν να έχουν ανθρώπινες προσωπικότητες, θα μπορούσε ίσως, να προκαλέσει σοβαρό ηθικό πρόβλημα. Μία μελέτη του 2013, από το πανεπιστήμιο της Ουάσιγκτον, έδειξε ότι ορισμένοι στρατιώτες που εργάζονταν παράλληλα με τα ρομπότ εξουδετέρωσης εκρηκτικών υλικών, συνδέθηκαν συναισθηματικά με αυτά και απελίζονταν όταν τα ρομπότ τους καταστρέφονταν. Ο κίνδυνος σύμφωνα με τον John P. Sullins, φιλόσοφο τεχνολογίας στο Sonoma State University, είναι ότι "η τάση μας να ανθρωπομορφοποιούμε τις μηχανές θα μπορούσε να μας κάνει ευάλωτους". Ο ίδιος πιστεύει ότι "δίνοντάς τους υψηλότερη ηθική ικανότητα απ' ότι στην πραγματικότητα διαθέτουν, δίνουμε συγχρόνως στις κερδοσκοπικές εταιρείες ρομποτικής, μια ευκαιρία να χειραγωγήσουν τους χρήστες με νέους κακόβουλους τρόπους".

Πώς θα πρέπει λοιπόν η ανθρωπότητα να αντιμετωπίσει τα ρομπότ; Ως αντικείμενα ή ως όντα; Κάθε ειδικός ρομποτικής έχει μία διαφορετική απάντηση. Πλέον, "μιλάμε για μία νέα κατηγορία που πρόκειται να συμπεριλάβει ένα ευρύ φάσμα τεχνολογιών", όπως ισχυρίζεται ο Sullins. Πιστεύει μάλιστα, ότι η άφιξη των νέων αυτών ρομποτικών όντων πρόκειται να ταράξει την ηθική. «Για χιλιάδες έτη κατά κάποιο τρόπο υπνοβατούσαμε εξηγώντας τι σημαίνει ηθική και δεοντολογία επειδή υποθέταμε ότι ο κόσμος ήταν όλος για μας, όλος για τις

ανθρώπινες σχέσεις» είπε. Τώρα πια βέβαια, ο σύγχρονος κόσμος, θέτει όλες αυτές τις έννοιες υπό βαθιά αμφισβήτηση.

Στις ΗΠΑ, η τάση αντιμετώπισης και αποδοχής των ρομπότ σαν ισάξια ή ακόμα και ανώτερα από τους ανθρώπους, έχει πάρει διαστάσεις. Σύμφωνα με το πολιτικό περιοδικό «Politico», ομάδα επιστημόνων και διανοουμένων, έφτασε στο σημείο να προτείνει να οριστεί ως Πρόεδρος της χώρας ένας υπολογιστής. Η ιδέα τους στηρίζεται στην σκέψη πως η τεχνητή νοημοσύνη θα κάνει την Αμερική μεγάλη και θα λαμβάνει αποφάσεις σε δύσκολα και πολύπλοκα ζητήματα, σωστότερα από οποιονδήποτε άνθρωπο. Παραδέχονται ωστόσο, πως το εγχείρημα αυτό δεν είναι καθόλου εύκολο, καθώς υπάρχουν πολλά εμπόδια.

Επίσης, ένας από τους οπαδούς του τρανσάνθρωπισμού (αναλυκότερη αναφορά του όρου στο κεφάλαιο 6.5), ο Zoltan Istvan, θεωρεί πως σύντομα η τεχνητή νοημοσύνη θα αναλάβει την εξουσία στις ΗΠΑ. «Ο Πρόεδρος-τεχνητή νοημοσύνη δεν θα μπορεί να εξαγοραστεί από λόμπι ούτε να επηρεαστεί από προσωπικές φιλοδοξίες και οικογενειακά κίνητρα. Δεν θα έχει τον νεποτισμό που βλέπουμε τώρα στον Λευκό Οίκο. Αυτά τα πράγματα δεν τα κάνουν οι μηχανές», τονίζει. Όλα αυτά θα επηρεάσουν σημαντικά τη ζωή των ανθρώπων και τον τρόπο σκέψης τους, άλλοτε θετικά και άλλοτε αρνητικά. Όλα αυτά μπορεί να ακούγονται από πολλούς ως σενάρια επιστημονικής φαντασίας αλλά τεχνολογικά είμαστε πολύ κοντά στην πραγμάτωσή τους.¹²⁶

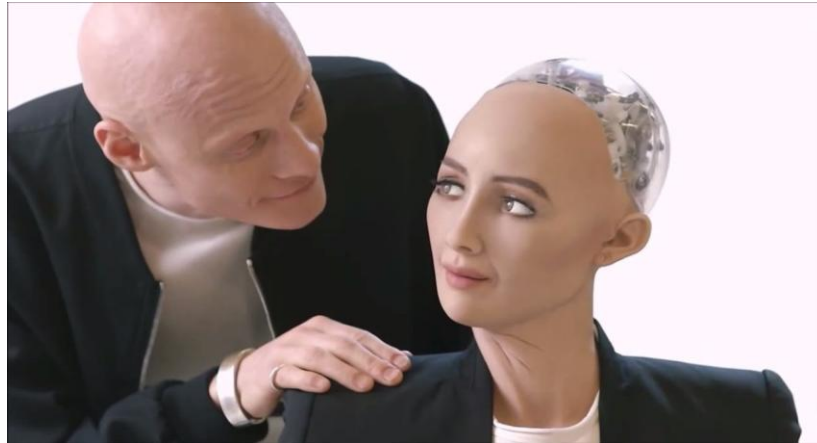
Αρκετοί επιστήμονες εκφράζουν προβληματισμούς, σχετικά με το πόσο παρεμβατική μπορεί να γίνει η τεχνολογία όσον αφορά στην καθημερινή ζωή και τα δικαιώματα των πολιτών. Η Google για παράδειγμα, για να καθησυχάσει τους προβληματισμούς αυτών των επιστημόνων, λέει τώρα ότι έχει στήσει ένα συμβούλιο ηθικής για να επιβλέπει την δουλειά που γίνεται στον συγκεκριμένο τομέα. Αλλά πώς γίνεται το ίδιο το συμβούλιο ηθικής της εταιρείας να βάζει φρένο στις επιχειρήσεις της, στις οποίες έχει μάλιστα, επενδύσει τεράστια κεφάλαια και με στόχο την κατασκευή λογισμικού που θα κάνει τις μηχανές να σκέφτονται ως άνθρωποι; Και αν ένα ρομπότ ή μια μηχανή μπορεί να σκέφτεται ως άνθρωπος και να λαμβάνει αποφάσεις, ποιος εγγυάται για τις σωστές του επιλογές; Η είδηση ότι τον ΟΗΕ τον απασχολεί ανεπίσημως έντονα το ζήτημα των αυτόνομων ρομπότ με δυνατότητα άσκησης θανάσιμης βίας, φέρνει ανησυχία.

Ακόμα κι αν λαμβάνονται όλες οι προδιαγραφές ασφαλείας και καθορίζονται πλήρως τι δικαιώματα θα δοθούν στις μηχανές, όπως λέει μια μερίδα επιστημόνων, ποιος μπορεί να αποκλείσει τον αστάθμητο ανθρώπινο παράγοντα, για παράδειγμα την οποιαδήποτε εταιρεία υπολογιστικής νοημοσύνης, που μπορεί να οραματίζεται την επικράτηση της δικής της αποκλειστικά εικονικής πραγματικότητας;¹²⁷

Ένα παράδειγμα για το πού μπορεί να φτάσει η τεχνητή και κατά επέκταση υπολογιστική νοημοσύνη, πόσο παρεμβατική μπορεί να γίνει στην καθημερινή ζωή και τι ηθικά διλλήματα μπορεί να δημιουργήσει, είναι το ανθρωποειδές ρομπότ, Σοφία. Η Σοφία, έχει σχεδιαστεί από την "Hanson Robotics", εταιρία του Χονγκ Κονγκ, βάσει της εικόνας της Όντρεϊ Χέμπορν. Συγκλονιστικό είναι το βίντεο στο οποίο η Σοφία, σε ρόλο ομιλήτη απαντά σε ερωτήσεις. "Από την μια συνειδητοποιούμε πόσο μπορεί να έχει προχωρήσει η τεχνολογία και από την άλλη ένας τρόμος διαπερνά το μυαλό μας", σχολιάζει το CNN. Η Σοφία μίλησε σε εκδήλωση στο Ριάντ της Σαουδικής Αραβίας, προκειμένου να εκφράσει τις ευχαριστίες της για την ιθαγένεια που της χορήγησε η χώρα, μια κίνηση την οποία περιέγραψε ως ιστορική. Πρόκειται για την πρώτη χώρα που χορήγησε υπηκοότητα σε ένα ανθρωποειδές ρομπότ.

¹²⁶ Κίμωνας Χαλαλάμπους, «Πρόταση σοκ για τον πλανήτη – ποιον θέλουν να φέρουν για να κυβερνήσει», Ιούλιος 2017

¹²⁷ Άννα Ρέμπελου, «Google: επικίνδυνα παιχνίδια με ρομπότ», Αύγουστος 2015



Εικόνα 18:

Σοφία - το ρομπότ με υπηκοότητα από τη Σαουδική Αραβία.

Επίσης, επιτέθηκε έμμεσα στον Έλον Μασκ, διευθύνοντα σύμβουλο της SpaceX και της Tesla. Όταν ο δημοσιογράφος είπε ότι «όλοι θέλουμε να αποφύγουμε ένα κακό μέλλον», η Σοφία απάντησε: "Έχετε διαβάσει πάρα πολύ τον Έλον Μασκ και παρακολουθείτε πάρα πολλές ταινίες του Χόλυγουντ. Μην ανησυχείτε, αν μου φέρεστε καλά, θα φέρομαι και εγώ καλά σε σας". Και το ρομπότ συνέχισε: "Αντιμετωπίστε με ως ένα έξυπνο σύστημα εξόδου-εισόδου". Ωστόσο, αξίζει να σημειωθεί ότι πολλοί είναι αντίθετοι στην απόφαση της Σαουδικής Αραβίας να δοθεί υπηκοότητα σε ένα ρομπότ. Ενδεικτικό είναι πως κάποιοι περιγράφουν την Σοφία ως "παράξενη" και άλλοι μιλούν για τον "Terminator".¹²⁸

Αλγόριθμοι και ρομπότ έχουν ξεκινήσει να κινούνται γύρω μας, ποιος όμως είναι υπεύθυνος για τις πράξεις τους; Μπορεί ένα ρομπότ ή ένα κομμάτι λογισμικού να φυλακιστεί αν διαπράξει ένα έγκλημα; Πώς θα πρέπει να κινηθούμε νομικά εάν ο κώδικας είναι εγκληματικός είτε από τον σχεδιαστή του είτε από προεπιλογή; Τι θα συμβεί εάν ένα ρομπότ ξεκινήσει να αγοράζει ναρκωτικά, όπλα, ή παράνομους εξοπλισμούς και αρχίσει να τους στέλνει σε εμάς και η αστυνομία παρακολουθεί το πακέτο; Όλα αυτά τα ερωτήματα είναι μέρος ενός φιλοσοφικού διλήμματος που προκύπτει λόγω της σκοτεινής πλευράς του διαδικτύου, εννοώντας την εμφάνιση των «darknet markets», την ανώνυμη χρήση του διαδικτύου και τα «bitcoins». Όλα αυτά τα ερωτήματα διερευνώνται σε μία νέα έκθεση που ονομάζεται «The darknet» και πραγματοποιείται στην Ελβετία.

Ένα επίσης δυνατό παράδειγμα είναι ένας αλγόριθμος ονομαζόμενος «Random Darknet Shopper», που δημιουργήθηκε το 2014, με σκοπό να διερευνήσει την σκοτεινή πλευρά του διαδικτύου. Το Random Darknet Shopper ήταν ένα αυτοματοποιημένο ρομπότ που πραγματοποιούσε ηλεκτρονικές αγορές σε bitcoins κάθε εβδομάδα. Μία φορά την εβδομάδα το ρομπότ έκανε τυχαίες αγορές μέσω του διαδικτύου και αποστέλλε απευθείας αυτές τις αγορές στον εκθεσιακό χώρο. Στην πρώτη εκτέλεσή του, από τον Οκτώβριο του 2014 έως τον Ιανουάριο του 2015, αγόρασε 12 τυχαία αντικείμενα από την παράνομη ιστοσελίδα «Agora». Τον Σεπτέμβριο αυτή η ιστοσελίδα διέκοψε τις υπηρεσίες της και ο αλγόριθμος άρχισε να παραγγέλνει αντικείμενα από το «Alpha Bay», μία από τις μεγαλύτερες ιστοσελίδες για ηλεκτρονικές αγορές που βρίσκεται στο «Deep Web». Ως "Deep Web" θεωρείται ο τόπος που οτιδήποτε είναι δυνατόν. Εκεί βρίσκεται κάθε είδος υλικών και υπηρεσιών για πώληση, και τα περισσότερα από αυτά είναι παράνομα. Θεωρείται ως έγκλημα στον κυβερνοχώρο αφού εκεί βρίσκονται υπηρεσίες για hacking, κλεμμένες πιστωτικές κάρτες, πορνογραφικό υλικό και πολλά άλλα. Μέσα σε αυτά που αγόρασε ο

¹²⁸ Newsroom, CNN Greece, «Σοφία το ρομπότ που πήρε υπηκοότητα από την Σαουδική Αραβία», Οκτώβριος 2017

αλγόριθμος, περιλαμβάνεται ένα καπέλο μπέιζμπολ με μία κρυμμένη κάμερα και 10 χαπάκια του ναρκωτικού έκστασης.

Η ελβετική αστυνομία "συνέλαβε" τον αλγόριθμο, κατάσχεσαν τον υπολογιστή καθώς και τα αντικείμενα που είχε αγοράσει το ρομπότ. Τα χαπάκια υποβλήθηκαν σε ιατροδικαστική εξέταση όπου και αποδείχθηκε ότι ήταν πραγματικά χαπάκια έκστασης. Τρεις μήνες αργότερα, το "έργο τέχνης" όπως το αποκαλούν οι δημιουργοί του, Carmen Weisskopf και Domagoj Smoljic, και όλα τα αντικείμενα που είχε αγοράσει αφέθηκαν ελεύθεροι, ενώ η έκσταση καταστράφηκε. Η σύλληψή του όμως, δημιουργεί σκέψεις όσον αφορά τα ακριβή όρια που χαράσσεται η διαχωριστική γραμμή μεταξύ ανθρώπινης και υπολογιστικής ενοχής. Σε περίπτωση που ένα ρομπότ αγοράσει μέσω του διαδικτύου χάπια θα πρέπει να υπόκειται στους ίδιους νόμους όπως ένας άνθρωπος έμπορος ναρκωτικών; Πώς θέλει η ανθρωπότητα να είναι ο ψηφιακός της κόσμος στο μέλλον; Ποιος είναι υπεύθυνος για την εγκληματική δράση ενός ρομπότ; Ο κατασκευαστής; Ο ιδιοκτήτης, ή το ίδιο το ρομπότ τελικά;

Σύμφωνα με τον δρ Δασκαλάκη, η Τεχνητή Νοημοσύνη μοιάζει με μωρό που αν και γεννιέται με κάποια χαρακτηριστικά, κατά βάθος είναι «tabula rasa» (άγραφος πίνακας). Παίρνει ερεθίσματα από τον γονέα, και σύμφωνα με αυτά σχηματίζει τις δικές του θέσεις. Κάτι ανάλογο συμβαίνει και με την ΑΙ, αφού η μάθησή της προέρχεται καθαρά από την διεπαφή με τον άνθρωπο. Απόδειξη αυτού είναι το συμβάν που δημιουργήθηκε με ένα, «chat bot», ρομπότ που αλληλοεπιδράει με τους χρήστες, καθώς χρήστες του επιτέθηκαν με ρατσιστικά και συνωμοσιολογικά σχόλια. Ο ίδιος ο καθηγητής είπε: "Μέσα σε 17 ώρες έγινε τρελός ρατσιστής και συνωμοσιολόγος". Άλλο ένα ζήτημα είναι εάν τα δεδομένα είναι επαρκή ώστε η ΑΙ να ενστερνιστεί στατιστικές που αντιπροσωπεύουν το δίκαιο. Γι αυτόν τον λόγο ο ίδιος αναφέρει ότι "Πρέπει να προστατέψουμε την τεχνητή νοημοσύνη από το να κάνει στατιστικά λάθη, αλλά το πρόβλημα είναι ότι η στατιστική είναι δύσκολη επιστήμη" σημειώσε.¹²⁹

Ένα ακόμα ηθικό δίλλημα είναι το εξής: "Σκεφτείτε ότι φτιάχνουμε αυτοοδηγούμενα αυτοκίνητα που κινούνται μαζικά στους δρόμους. Αναπόφευκτα κάποιο από αυτά θα βρει τον εαυτό του σε φάση αναγνώρισης του γεγονότος ότι σε μερικά δευτερόλεπτα θα γίνει ένα αναπόφευκτο ατύχημα με εμπλοκή πεζών. Ο αλγόριθμος που οδηγεί καταλαβαίνει τότε ότι έχει δύο δυνατότητες: να πάει ευθεία και να σκοτώσει τους πεζούς ή να πάει αριστερά, να χτυπήσει στο στηθαίο και να σκοτώσει τους επιβαίνοντες. Δεν μπορεί να σώσει και τους δύο. Πώς θα πάρει την απόφαση; Ο αλγόριθμος μπορεί επίσης να καταλαβαίνει ότι οι πεζοί είναι ένα παιδάκι 8 χρονών, ο μπαμπάς του 41, και ο σκύλος τους και οι επιβαίνοντες μια έγκυος γυναίκα 30 ετών και το αγοράκι της. Πώς εγώ που σχεδιάζω τον αλγόριθμο θα λάβω την απόφαση για το ποιος θα ζήσει";¹³⁰

6.4 SUPERINTELLIGENCE και ΜΟΝΑΔΙΚΟΤΗΤΑ - Ηθικά διλήμματα που προκύπτουν.

Τα ηθικά ζητήματα που σχετίζονται με την πιθανή μελλοντική δημιουργία των μηχανών με γενικές πνευματικές ικανότητες, ξεπερνούν κατά πολύ εκείνα των ανθρώπων και είναι αρκετά διαφορετικά από οποιαδήποτε ηθικά προβλήματα που προκύπτουν στα τρέχοντα συστήματα αυτοματισμού και πληροφοριών. Μία τέτοια εξαιρετική νοημοσύνη δεν θα είναι απλώς μια άλλη τεχνολογική ανάπτυξη, θα είναι η πιο σημαντική εφεύρεση που έγινε ποτέ. Θα οδηγήσει σε εκρηκτική πρόοδο όλους τους επιστημονικούς και τεχνολογικούς τομείς,

¹²⁹ ΑΠΕ-ΜΠΕ, «Η συγκλονιστική διάλεξη του 37χρονου καθηγητή ιδιοφυΐα Κωνσταντίνου Δασκαλάκη στο ΑΠΘ», Ιανουάριος 2018

¹³⁰ Zimzamphysics.gr, 6 Φεβρουαρίου 2018

καθώς μια *superintelligence* θα διεξάγει έρευνα με υπεράνθρωπη αποτελεσματικότητα. Στο βαθμό που η ηθική είναι μια γνωστική άσκηση, μία τέτοιου είδους νοημοσύνη θα μπορούσε επίσης να ξεπεράσει εύκολα τους ανθρώπους στην ποιότητα της ηθικής σκέψης τους. Ωστόσο, οι σχεδιαστές της *superintelligence* θα πρέπει να καθορίσουν τα αρχικά κίνητρά της. Δεδομένου ότι μπορεί να γίνει ασταμάτητα ισχυρή, λόγω της πνευματικής ανωτερότητάς της και των τεχνολογιών που θα μπορούσε να αναπτύξει, είναι σημαντικό να παρέχεται με κίνητρα φιλικά προς τον άνθρωπο.

Μέσα από αυτές τις προκλήσεις δημιουργούνται πολλά φιλοσοφικά και ηθικά ερωτήματα. Αν υποθέσουμε πως δημιουργείται ένας υπολογιστής ο οποίος μπορεί να μιλήσει και να ενεργήσει ανθρώπινα, τότε μπορεί να αισθανθεί όπως ένα ανθρώπινο ον; Προχωρώντας βαθύτερα δημιουργούνται κι άλλα σημαντικά ερωτήματα. Εάν μπορεί να μεταφερθεί η συνείδηση ενός ανθρώπου πριν να πεθάνει σε έναν υπολογιστή, θα υπάρχει αυτός ο άνθρωπος ακόμα κι αν έχει πεθάνει; Από ποιον θα κρίνεται το δικαίωμα κάποιου στην αθανασία; Με ποια κριτήρια θα αποφασίζεται εάν ένα ον έχει αισθήματα; Ενώ η ανθρώπινη ζωή πλησιάζει την αθανασία, θα εξακολουθήσει να έχει το ίδιο νόημα; Αν νικήσει η ανθρωπότητα τον θάνατο, μήπως χάσει τελικά την ανθρωπιά της;¹³¹

Είναι δύσκολο να υπάρξει κάποιο πρόβλημα που ένα *superintelligence* δεν θα ήταν σε θέση να το λύσει ή τουλάχιστον να βοηθήσει στην επίλυσή του. Η ασθένεια, η φτώχεια, η καταστροφή του περιβάλλοντος, η άσκοπη ταλαιπωρία όλων των ειδών, είναι κάποιες από τις άσχημες καταστάσεις που μια *superintelligence*, εξοπλισμένο με προηγμένη νανοτεχνολογία, θα είναι ικανό να εξαλείψει. Επιπλέον, θα μπορούσε να δώσει αόριστη διάρκεια ζωής στον άνθρωπο, είτε σταματώντας και αντιστρέφοντας τη διαδικασία της γήρανσης μέσω της χρήσης της νανοϊατρικής¹³², ή προσφέροντάς του τη δυνατότητα να αναβαθμίσει ο ίδιος τον εαυτό του. Ένα *superintelligence* θα μπορούσε επίσης να δημιουργήσει πολλές ευκαιρίες στον άνθρωπο, όπως να αυξήσει κατά πολύ τις πνευματικές και συναισθηματικές του ικανότητες, και θα μπορούσε να βοηθήσει στη δημιουργία ενός άκρως ελκυστικού βιοματικού κόσμου στον οποίο θα μπορούσε να ζήσει μία χαρούμενη και ιδανική ζωή.

Ωστόσο, τόσο λόγω της ανώτερης ικανότητας σχεδιασμού όσο και λόγω των τεχνολογιών που θα μπορούσαν να αναπτυχθούν, είναι πολύ πιθανό, η πρώτη *superintelligence* να είναι πολύ ισχυρή και ασυναγώνιστη. Θα είναι σε θέση να επιφέρει σχεδόν κάθε πιθανό δυνατό αποτέλεσμα και να εμποδίσει κάθε προσπάθεια που τυχόν θα αποτρέψει την εφαρμογή των κορυφαίων στόχων της. Θα μπορούσε να εξοντώσει όλους τους άλλους παράγοντες, να τους πείσει να αλλάξουν τη συμπεριφορά τους, ή να εμποδίσει τις προσπάθειές τους σε παρεμβολές. Ακόμη και ένα «*περιορισμένο superintelligence*» (*fettered superintelligence*) που εκτελούνταν σε έναν απομονωμένο υπολογιστή, ικανό να αλληλοεπιδράσει με τον υπόλοιπο κόσμο μόνο μέσω της διεπαφής κειμένων, ίσως να είναι σε θέση να ξεφύγει από τον περιορισμό του, πείθοντας τους χειριστές του να το απελευθερώσουν. Υπάρχουν ακόμη και κάποια προκαταρκτικά πειραματικά αποδεικτικά στοιχεία, ότι αυτό θα μπορούσε να συμβεί.¹³³

Φαίνεται ότι ο καλύτερος τρόπος για να εξασφαλιστεί ότι μία *superintelligence* θα έχει ευεργετική επίδραση στον κόσμο είναι να εφοδιαστεί με φιλανθρωπικές αξίες. Ο κορυφαίος στόχος τους πρέπει να είναι η φιλικότητα.¹³⁴ Εάν όμως, μια *superintelligence* ξεκινήσει έχοντας σαν κορυφαίο στόχο της να είναι φιλική, τότε μπορεί να φροντίσει να παραμείνει φιλική, ή τουλάχιστον να μην απαλλαγεί συνειδητά από την φιλικότητά της; Το σημείο αυτό είναι στοιχειώδες. Πώς ακριβώς η φιλικότητα πρέπει να γίνει κατανοητή και πώς πρέπει να

¹³¹ Ομιλία του RAY KURZWEIL, «Ένα πανεπιστήμιο για την επερχόμενη μοναδικότητα», TED 2009

¹³² Freitas Jr., R. A *Nanomedicine, Volume 1: «Basic Capabilities»*, 1999

¹³³ Yudkowsky, «*The AI Box Experiment*», 2002

¹³⁴ Yudkowsky, "*Creating Friendly AI 1.0.*", 2003

εφαρμοστεί, και πώς η ομόνοια θα πρέπει να κατανεμηθεί μεταξύ των διαφορετικών ανθρώπων και των μη ανθρωπίνων πλασμάτων, είναι ένα θέμα που χρήζει περαιτέρω εξέτασης.

Επίσης, ένα άλλο σημαντικό ζήτημα είναι, αν όλοι οι άνθρωποι, αλλά και άλλα αισθανόμενα πλάσματα στη γη, θα μπορούσαν να πάρουν ένα σημαντικό μερίδιο της ευεργεσίας μιας *superintelligence*. Εάν τα οφέλη που θα μπορούσε να παραχωρήσει είναι απέραντα, τότε το σημαντικότερο είναι να εξασφαλιστεί ότι όλοι θα παίρνουν κάποιο μερίδιο, τέτοιο ώστε, να μπορεί να εγγυηθεί μία πολυετή και καλή ζωή για όλους.

Ένας άλλος κορυφαίος στόχος πρέπει να είναι η φιλανθρωπία. Ένας από τους κινδύνους που ελλοχεύει ωστόσο αν της δοθεί, είναι ότι εκείνοι που αναπτύσσουν τη *superintelligence* ίσως να μην την καταστήσουν γενικά φιλανθρωπική, αντίθετα να της δώσουν έναν πιο περιορισμένο στόχο εξυπηρετώντας μόνο τα συμφέροντα κάποιας μικρής ομάδας, όπως των ίδιων των δημιουργών της ή εκείνων που τους ανατέθηκε.

Ένας άλλος κίνδυνος επίσης είναι, εάν κατά λάθος μια καλοπροαίρετη ομάδα προγραμματιστών κάνει ένα μεγάλο λάθος στο σχεδιασμό της και οδηγήσει σε μία *superintelligence* της οποίας ο κορυφαίος στόχος της είναι η κατασκευή ενός προϊόντος, παραδείγματος χάριν συνδετήρων όπως αναφέρθηκε σαν παράδειγμα σε προηγούμενο κεφάλαιο, θα μπορούσε να μετατρέψει ολόκληρη τη γη σε εγκαταστάσεις παραγωγής συνδετήρα. Με άλλα λόγια, ίσως η δημιουργία μιας *superintelligence* ακόμα κι αν πραγματοποιήσει τον επιθυμητό σχεδιασμό, στην πραγματικότητα να αποδειχτεί ότι είναι μια ψεύτικη ουτοπία, στην οποία τα πράγματα ουσιαστικής σημασίας για την ανθρώπινη άνθηση θα έχουν χαθεί μία για πάντα ανεπιστρεπτί.

Από την άλλη, ο Nick Bostrom επίσης αναφέρει, ότι στο βαθμό που η ηθική είναι μια γνωστική επίτευξη, μια *superintelligence* θα μπορούσε να είναι πολύ καλύτερη σε αυτό από ότι ένας ανθρώπινος φιλόσοφος. Αυτό σημαίνει ότι οι ερωτήσεις σχετικά με την ηθική, στο μέτρο που αυτές έχουν σωστές απαντήσεις οι οποίες μπορούν να προκύψουν από το συλλογισμό και την στάθμιση των αποδεικτικών στοιχείων, θα μπορούσαν να απαντηθούν με μεγαλύτερη ακρίβεια από μία *superintelligence* απ' ότι από τους ανθρώπους. Το ίδιο ισχύει και για τα ζητήματα πολιτικής και του μακροπρόθεσμου προγραμματισμού, όσο αφορά την κατανόηση ποιες πολιτικές θα οδηγήσουν σε ποια αποτελέσματα, οπότε θα είναι πιο αποτελεσματικό όσον αφορά την επίτευξη των δεδομένων στόχων. Μια *superintelligence* μπορεί να ξεπεράσει τους ανθρώπους.

Υπάρχουν επομένως, πολλές ερωτήσεις που μια *superintelligence* θα μπορούσε να απαντήσει καλύτερα και γρηγορότερα από τον άνθρωπο. Έχει τη δυνατότητα να εκτιμήσει και να αξιολογήσει τα δεδομένα χωρίς να σπαταλήσει πολύ χρόνο, έχει περισσότερη μνήμη και καλύτερη νοημοσύνη και πληροφόρηση. Ο Nick Bostrom ισχυρίζεται επίσης, ότι θα μπορούσε να της ζητηθεί να βοηθήσει τον προγραμματιστή της να προσδιορίσει την πραγματική πρόθεση του αιτήματός του, όταν διαμορφώνει ένα στόχο γι' αυτήν, ώστε να μειωθεί ο κίνδυνος κάποια ατυχής διατύπωση ή σύγχυση σχετικά με το τι θέλει να πετύχει, να οδηγήσει σε ανεπιθύμητα αποτελέσματα.¹³⁵

Ένα σημαντικό ζήτημα που θα πρέπει να ληφθεί υπόψη πριν την ανάπτυξη μιας *superintelligence* τεχνολογίας, είναι η χρονική στιγμή που θα επιλεγεί για να αναπτυχθεί. Πρέπει να εξεταστεί από τους επιστήμονες, εάν θα ήταν καλύτερο να επιταχυνθεί η ανάπτυξη της ή να καθυστερήσει. Μια *superintelligence* θα μπορούσε να βοηθήσει να μειωθούν ή να

¹³⁵ Nick Bostrom, «Ethical Issues in Advances Artificial Intelligence», pp.12-16, Πανεπιστήμιο Oxford, 2003

εξαλειφθούν άλλοι υπαρξιακοί κίνδυνοι,¹³⁶ όπως για παράδειγμα ο κίνδυνος ότι η προηγμένη νανοτεχνολογία θα χρησιμοποιηθεί από τους ανθρώπους για εχθροπραξίες ή για την τρομοκρατία, μια σοβαρή απειλή για τη μακροπρόθεσμη επιβίωση της ευφυούς ζωής στη γη. Ο Nick Bostrom ισχυρίζεται ότι "αν έχουμε την ευκαιρία να φτάσουμε πρώτα στη *superintelligence*, μπορούμε πιθανότατα να αποφύγουμε αυτόν τον κίνδυνο που αφορά τη νανοτεχνολογία και πολλά άλλα. Εάν, από την άλλη πλευρά, φτάσουμε πρώτα στη νανοτεχνολογία, θα πρέπει να αντιμετωπίσουμε και τους κινδύνους που θα προκληθούν με την χρήση της και, κατά συνέπεια, τους κινδύνους από την *superintelligence*".

Ο συνολικός κίνδυνος φαίνεται να ελαχιστοποιείται εφαρμόζοντας την *superintelligence*, με μεγάλη προσοχή, όσο το συντομότερο δυνατόν. Σύμφωνα με τον Nick Bostrom, δεν μπορεί η επιστημονική κοινότητα να μείνει αδιάφορη για το πώς θα κατασκευαστεί. Αντιθέτως, η συγκρότηση και η δημιουργία των αρχικών συνθηκών, και κυρίως η επιλογή ενός στόχου υψηλού επιπέδου για μία *superintelligence*, είναι υψίστης σημασίας. "Ολόκληρο το μέλλον μας μπορεί να εξαρτηθεί από το πώς θα λύσουμε αυτά τα προβλήματα", είπε.

Επιπλέον, είναι πολύ πιθανό, μια *Superintelligence* να οδηγήσει σε μία πιο προχωρημένη *superintelligence*. Αυτό θα προκύψει τόσο από το βελτιωμένο υλικό που θα μπορούσε να δημιουργήσει, αλλά και από τις βελτιώσεις που θα μπορούσε να κάνει στον δικό της πηγαίο κώδικα. Επίσης, από την στιγμή που οι τεχνητές νοημοσύνες είναι λογισμικό, μπορούν εύκολα και γρήγορα να αντιγραφούν, εφ' όσον υπάρχει διαθέσιμο hardware για να αποθηκευτούν. Το hardware και το κόστος της δημιουργίας ενός επιπλέον αντιγράφου μίας τεχνητής νοημοσύνης είναι κοντά στο μηδέν. Άρα τα τεχνητά μυαλά θα μπορούσαν να υπάρχουν σε μεγάλους αριθμούς, παρόλο που είναι πιθανό η αποτελεσματικότητα να ευνοήσει τη συγκέντρωση υπολογιστικών πόρων σε μια ενιαία υπερ-διάνεια.

Μια *superintelligence* δεν θα έπρεπε απαραίτητως να θεωρηθεί ως ένα απλό εργαλείο. Ενώ τα εξειδικευμένα *superintelligences* που μπορούν να σκέφτονται μόνο για ένα περιορισμένο σύνολο θεμάτων μπορούν να είναι εφικτά, ένα γενικό *superintelligence* θα ήταν ικανό να παίρνει ανεξάρτητες πρωτοβουλίες και να παράγει τα δικά του σχέδια. Ως εκ τούτου, θα μπορεί να σκεφτεί πιο κατάλληλα ως δυνητικά αυτόνομος παράγοντας. Ενώ μπορεί να διαρκέσει πολύ μέχρι να φτάσουμε στην *superintelligence*, το τελικό στάδιο μπορεί να συμβεί γρήγορα. Η μετάβαση από μια κατάσταση όπου υπάρχει κατά προσέγγιση ένα ανθρώπινο επίπεδο τεχνητής νοημοσύνης σε μια κατάσταση όπου θα υπάρξει πλήρης άνθηση της *superintelligence*, με επαναστατικές εφαρμογές, μπορεί να είναι πολύ γρήγορη, σημειώνει ο Nick Bostrom.

Ενώ οι άνθρωποι δεν αρέσκονται στο να γίνονται σκλάβοι, μία *superintelligence* μπορεί να κατασκευαστεί με τέτοιο τρόπο ώστε, να έχει ως στόχο ύπαρξής της, να υπηρετήσει την ανθρωπότητα ή κάποιο συγκεκριμένο άνθρωπο, χωρίς καμία απολύτως επιθυμία για να εξεγερθεί, να επαναστατήσει ή να "απελευθερώσει" τον εαυτό της. Φαίνεται, ωστόσο, απόλυτα πιθανό να υπάρξει μία *superintelligence* της οποίας μοναδικός της στόχος θα είναι κάτι εντελώς αυθαίρετο, το οποίο θα υποστήριζε με όλες της τις δυνάμεις.

Οι τεχνητές διάνοιες πιθανότατα θα είναι εύκολο να προφυλαχθούν από ορισμένα είδη ανθρώπινου λάθους και προκατάληψης, ενώ την ίδια στιγμή διατρέχουν αυξημένο κίνδυνο από άλλα είδη λάθους που ακόμη και ο πιο απρόσεκτος άνθρωπος δεν θα έκανε. Υποκειμενικά, η πιθανή ενδότερη συνειδητή ζωή μιας τεχνητής νοημοσύνης, κάτι που

¹³⁶ Bostrom, "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards" *Journal of Evolution and Technology*, 2002

να μοιάζει με την ανθρώπινη ψυχή, μπορεί να είναι πολύ διαφορετική από αυτήν ή ακόμα, να μην μοιάζει καθόλου.

Για όλους αυτούς τους λόγους, ο Nick Bostrom θεωρεί ότι, πρέπει κάποιος να είναι δύσπιστος όσον αφορά την υπόθεση ότι η εμφάνιση της superintelligence μπορεί να προβλεφθεί προεκτείνοντας την ιστορία άλλων τεχνολογικών ανακαλύψεων, ή ότι η φύση και οι συμπεριφορές των τεχνητών διανοιών θα μοιάζουν απαραίτητα με εκείνες του ανθρώπινου ή άλλων ζωικών μυαλών. Εν κατακλείδι αναρωτιέται: *"Το ανθρώπινο είδος επικράτησε των άλλων ειδών εξαιτίας της νοήμονος συμπεριφοράς του. Τι θα συμβεί όμως αν δημιουργηθεί μια τεχνητή νοημοσύνη η οποία θα υπερβαίνει αυτήν του ανθρώπου;"*

Ο Ray Kurzweil, συγγραφέας, εφευρέτης, ιδρυτής του πανεπιστημίου Singularity και υπερ-εγκέφαλος του MIT, υποστηρίζει ότι *«Μέχρι το τέλος της δεκαετίας, οι υπολογιστές θα είναι το ίδιο έξυπνοι με τον άνθρωπο. Μέχρι το 2045, θα γίνουμε αθάνατοι»*. Τότε τοποθετεί την ημερομηνία για την "Μοναδικότητα". Αυτό βέβαια δεν αναφέρεται στο ανθρώπινο σώμα αλλά στην νοημοσύνη του, που μπορεί να παραμείνει ζωντανή με την βοήθεια της ΑΙ.

Ο Kurzweil ισχυρίζεται ότι πλησιάζει η μέρα που τα υπολογιστικά συστήματα θα γίνουν όχι απλώς πιο έξυπνα, αλλά πιο έξυπνα από τα ανθρώπινα όντα. Αυτό θα έχει σοβαρό αντίκτυπο στην ανθρωπότητα, αφού θα την αλλοιώσει οριστικά. Με τους υπολογισμούς που έκανε, βρήκε ότι ο πολιτισμικός θάνατος της ανθρωπότητας θα συμβεί σε λιγότερο από 30 χρόνια. Κάτι που είναι όχι μόνο αναμενόμενο σύμφωνα με αυτόν, αλλά και αναπόφευκτο.

Αναμφισβήτητα, η υπολογιστική νοημοσύνη αναπτύσσεται ταχύτατα. Από την στιγμή που αυτά τα συστήματα θα γίνονται ολοένα και πιο γρήγορα, είναι πολύ πιθανό να κατακτήσουν το επίπεδο της ανθρώπινης νοημοσύνης. Η λειτουργία του ανθρώπινου εγκεφάλου που διαμορφώνει την συνείδηση, θα μπορούσε εύκολα να εξομοιωθεί από αυτήν την υπερδύναμη. Έτσι η χρήση της δεν θα είναι μόνο για γρήγορες αριθμητικές πράξεις, και σύνθεση μουσικής αλλά και για να λαμβάνει ηθικές αποφάσεις και εκτιμήσεις.

Ο ίδιος θεωρεί ότι δεν θα μπορέσει κανένας να σταματήσει την εξέλιξη των υπολογιστικών συστημάτων, τα οποία θα εξελίσσονται μέχρι να αναπτύξουν πολύ μεγαλύτερη ευφυΐα από την ανθρώπινη. Θα φτάσουν σε ένα επίπεδο που η ευφυΐα τους θα κατακτά ολοένα και μεγαλύτερα επίπεδα, αφού πλέον η εξέλιξή τους θα εξαρτάται από τα ίδια και όχι από τον δημιουργό τους. *"Φανταστείτε έναν επιστήμονα πληροφορικής που έχει τις δυνατότητες ενός υπερ-ευφυή υπολογιστή. Θα έκανε την δουλειά απίστευτα γρήγορα. Θα μπορούσε να αντλεί πληροφόρηση από τεράστιες ποσότητες δεδομένων χωρίς κόπο. Δεν θα χρειαζόταν καν διαλείμματα"*, καταλήγει.

Είναι αδύνατο να προβλέψει ο άνθρωπος πώς θα συμπεριφερθούν οντότητες πιο ευφυείς από τον ίδιο, στην περίπτωση που χρειαστεί να συνυπάρξουν στον κόσμο. Υπάρχουν πολλές θεωρίες σχετικά με την προσπάθεια του ανθρώπου, αλλά και τη δυνατότητά του να γίνει το ίδιο έξυπνος με αυτές τις οντότητες. Κάποιες ισχυρίζονται ότι θα συγχωνευθεί με αυτές, ώστε να γίνει και ο ίδιος μία υπερδιάνοια, κάνοντας χρήση της υπολογιστικής δύναμης για να φτάσει σε υψηλότερα επίπεδα τις πνευματικές του ικανότητες. Άλλες θεωρίες, υποστηρίζουν ότι η ΑΙ θα βοηθήσει τον άνθρωπο να αντιμετωπίσει τα γηρατειά και να παρατείνει την ζωή του μέχρι την αθανασία. Άλλες, ότι θα μεταφέρει τη συνείδησή του στα υπολογιστικά συστήματα καταφέροντας έτσι να ζήσει μέσα στο λογισμικό τους επ' αόριστον. Υπάρχουν βέβαια κι αυτές που πρεσβεύουν, ότι οι υπολογιστές θα πάρουν ανθρώπινη μορφή και θα εξολοθρεύσουν την ανθρωπότητα.

Σύμφωνα με τον Scott Adams, *«Κάποια μέρα θα παρευρίσκεστε στην κηδεία του παππού σας και μία ώρα αργότερα η ρομποτική έκδοση του παππού σας θα βρίσκεται στην πόρτα και θα σας ρωτήσει πώς πήγε. Η ημέρα που θα χάσετε κάποιον θα είναι η ημέρα που η*

ρομποτική έκδοση θα έχει ενεργοποιηθεί νόμιμα. Δεν θα μπορείτε να έχετε δύο εκδόσεις του ίδιου ατόμου γιατί θα ήταν ένα χάος».

Ο ίδιος πιστεύει ότι στο μέλλον όλα τα ρομπότ θα είναι πιο ανεπτυγμένα από τις ίδιες τις προσωπικότητες των ανθρώπων που τα δημιούργησαν. Η ανθρωπότητα θα αφήσει πίσω την κατάθλιψη και τις κακές εμπειρίες. Θα ενισχύει την ευτυχία, την αντοχή προς τους άλλους, και ίσως την ίδια της την ακεραιότητα. Η ρομποτική έκδοση του εαυτού ενός ανθρώπου θα έχει καλύτερα χαρακτηριστικά από ότι ο ίδιος ο άνθρωπος. Επίσης ισχυρίζεται ότι: *«Την πρώτη μέρα, η ρομποτική σας μορφή θα μοιάζει με την παλιά σας ανθρώπινη μορφή με ότι ατέλεια είχατε. Αλλά με την πάροδο του χρόνου θα αναβαθμίσετε το ρομποτικό σας σώμα ώστε να είναι πιο ελκυστικό, γιατί η ομορφιά προσφέρει ένα πλεονέκτημα σε σχέση με την κοινωνία μας».*

Οι μελέτες που έκανε ο Kurzweil έδειξαν, όπως ισχυρίστηκε στην ομιλία του στο TED το 2009, ότι θα αποκωδικοποιηθεί ο ανθρώπινος εγκέφαλος και τα υπολογιστικά συστήματα θα αποκτήσουν ευφυΐα ισάξια της ανθρώπινης, μέχρι το τέλος της δεκαετίας του 2020. Επίσης, ανέφερε ότι με το να αυξηθεί τόσο ραγδαία η υπολογιστική ισχύ και να ελαττωθεί τόσο πολύ το κόστος της, οι τεχνητές νοημοσύνες που θα υπάρχουν στον κόσμο, θα είναι *«σχεδόν ένα δισεκατομμύριο φορές περισσότερες από όλη την ανθρώπινη νοημοσύνη που υπάρχει σήμερα».*

Πολλοί είναι αυτοί που δυσκολεύονται να πιστέψουν ότι η "Μοναδικότητα" δεν είναι κάτι φανταστικό και εξωπραγματικό. Το θέμα όμως αυτό, χρήζει σοβαρότητας και σκέψης, αφού το μέλλον του πλανήτη Γη θα εξαρτηθεί από αυτήν. Είναι δύσκολο κάποιος να αντιληφθεί και να συλλάβει αυτήν την σκέψη, σχετικά με τις αθάνατες υπερδιάνοιες. Γι αυτόν τον λόγο, επενδύονται πολλά χρήματα σε αυτόν τον τομέα. Πριν από λίγα χρόνια, δημιουργήθηκε το Πανεπιστήμιο της Μοναδικότητας, το οποίο και φιλοξενεί η NASA. Ακόμα και εάν η εκτίμηση, του τι πρόκειται να γίνει στο μέλλον, είναι σοκαριστική, η ανθρωπότητα ελκύεται από την σκέψη της "Μοναδικότητας". Στο ενδεχόμενο βέβαια να πραγματοποιηθεί στο μέλλον, θα είναι ό,τι πιο σημαντικό έχει συμβεί στην ανθρωπότητα, μετά την ανακάλυψη της γραφής από τον άνθρωπο.¹³⁷

6.5 Ηθικά ζητήματα που εγείρονται από τη χρήση ιατρικών υπολογιστικών συστημάτων.

Οι ρομποτικές τεχνολογίες, όπως αναφέρθηκε στο τέταρτο κεφάλαιο της παρούσης εργασίας, στην ενότητα 4.6, βοηθούν τους γιατρούς να κάνουν καλύτερες διαγνώσεις και να παρέχουν ανώτερης ποιότητας θεραπείες. Από χειρουργικές επεμβάσεις μέχρι και θεραπείες αποκατάστασης, τα ρομπότ έχουν αρχίσει ήδη να εκτελούν περίπλοκες διαδικασίες που, σε πολλές περιπτώσεις, δεν θα μπορούσαν να επιτευχθούν με τα "γυμνά" ανθρώπινα χέρια. Τα υπολογιστικά συστήματα είναι τα πλέον κατάλληλα για την πρόβλεψη μιας ιατρικής πάθησης, αλλά πολλές φορές και για την ίασή της. Τα ρομποτικά συστήματα επίσης, βοηθούν πολύ σε αυτό.

Η χρήση όμως των υπολογιστικών συστημάτων σε θεραπευτικά πλαίσια, εγείρει πολλά ηθικά ζητήματα. Η ανθρώπινη επαφή αποτελεί μια από τις πιο θεμελιώδεις πτυχές της ανθρώπινης φροντίδας. Η αντικατάσταση του ανθρώπινου παράγοντα από ρομπότ, θα αφαιρούσε το ανθρώπινο στοιχείο από τις πρακτικές φροντίδας. Ένα άλλο ηθικό ζήτημα που προκύπτει είναι, ο βαθμός αυτονομίας που πρέπει να δοθεί στα ρομπότ. Για παράδειγμα, οι γονείς θα ήταν ικανοποιημένοι να αφήσουν τη φροντίδα του παιδιού τους στα χέρια ενός ρομπότ; Η «Ηθική» θα πρέπει να αντιμετωπίσει αυτά τα ζητήματα διασφαλίζοντας ότι όλοι οι

¹³⁷ Ομιλία του RAY KURZWEIL στο TED «Ένα πανεπιστήμιο για την επερχόμενη μοναδικότητα», άρθρο «2015: Το έτος που ο άνθρωπος θα γίνει αθάνατος», Νοέμβριος 2011

γνωστοί ηθικοί περιορισμοί και οι υφιστάμενοι νόμοι που ισχύουν εφαρμόζονται στη διαδικασία σχεδιασμού, εξετάζοντας συγκεκριμένα ηθικά ζητήματα που προκύπτουν από την αλληλεπίδραση μεταξύ του ασθενούς και των αυτόνομων θεραπευτικών ρομπότ, και ερευνώντας εάν και πώς μπορούν να οικοδομηθούν ηθικοί περιορισμοί για αυτά.¹³⁸

Υπάρχει μια κίνηση που ξεκίνησε το 1960 και ονειρεύεται ένα μέλλον στο οποίο θα κυριαρχούν αθάνατοι υπεράνθρωποι με τη βοήθεια και τη δύναμη της επιστήμης και ονομάζεται τρανσάνθρωπισμός/διανθρωπισμός ή τρανσουμανισμός (Transhumanism). Παρουσιάζεται σαν μία πολιτιστική κίνηση, η οποία στηρίζει και προωθεί την χρήση της τεχνολογίας και της επιστήμης για την βελτίωση των φυσικών αλλά και πνευματικών χαρακτηριστικών του ανθρώπου. Την παρέμβαση δηλαδή ακόμα και στο πνεύμα, στο σύνολο των πληροφοριών που συνθέτουν τον τρόπο σκέψης, την προσωπικότητα και τον χαρακτήρα του ανθρώπου. Αυτό δημιουργεί πολλά ηθικά ζητήματα. Οι απόψεις των επιστημόνων δίστανται.

Σύμφωνα με τον Max Moore «Ο Τρανσουμανισμός είναι μία σειρά από Φιλοσοφίες της Ζωής που αναζητούν την συνέχιση και επιτάχυνση της εξέλιξης της ζωής πέρα από την ανθρώπινη μορφή της, αλλά και τους ανθρώπινους περιορισμούς, με επιστημονικά και τεχνολογικά μέσα και αρχές που προωθούν την ίδια τη ζωή».¹³⁹

Η ευφυΐα, το συναίσθημα και η ικανότητα των ανθρώπων να βάζουν στόχους, μπορούν να πολλαπλασιαστούν από τεχνολογίες όπως είναι οι: «συνεπαφές ανθρώπινου εγκεφάλου-υπολογιστή» και η «Νευροφαρμακολογία». Ο φιλόσοφος Keith A. Bauer, στην εργασία του που ονομάζεται "Wired Patients" εξηγεί τον τρόπο που ηλεκτρόδια μπορούν να συνδεθούν σε ανθρώπινους εγκεφάλους, βοηθώντας να ανακτηθούν λειτουργίες ασθενών, που έχουν απενεργοποιηθεί μετά από εγκεφαλικά επεισόδια ή τραυματισμούς στο ωτιαίο μυελό.

Στη Ιατρική, όπως εξηγεί ο Bauer, οι βιοανιχνευτές και τα microchip, μπορούν να ενισχύσουν την ιατρική περίθαλψη, αφού δίνουν την δυνατότητα στους ιατρούς να αντιληφθούν και να αποτρέψουν ένα συμβάν που μπορεί να έχει τραγική κατάληξη, πριν καν εμφανιστεί. Κάπως έτσι για παράδειγμα μπορεί να αποφευχθεί μια καρδιακή προσβολή. Όσο καλό και εάν ακούγεται, κάτι τέτοιο θα μπορούσε να διαταράξει τις σχέσεις μεταξύ ιατρών και ασθενών καθώς και να κινδυνεύει το «Ιατρικό Απόρρητο». Παλαιότερα, ό,τι είχε σχέση με τις προοπτικές της γενετικής, φαίνονταν αποκύημα φαντασίας. Στις μέρες μας όμως αποδείχτηκε ότι με αυτήν μπορεί να ενισχυθεί το ανθρώπινο είδος και τα σενάρια που τότε φαίνονταν αδύνατα, τώρα πλέον εμφανίζονται ως πολύ πιθανά.

Η επέμβαση της Υψηλής Τεχνολογίας στο ανθρώπινο σώμα μπορεί να βοηθήσει στην επιδιόρθωση των τραυματισμένων περιοχών, τα φάρμακα μπορούν να βοηθήσουν στην εγκεφαλική διαύγεια ενώ βιονικά εμφυτεύματα μπορούν να βελτιώσουν τη όραση και την κινητικότητα του ανθρώπου. Οδεύουμε όμως προς ένα μέλλον το οποίο θα περιλαμβάνει ανθρώπους οι οποίοι θα είναι γενετικά σχεδιασμένοι ώστε: να έχουν εξαιρετική ευφυΐα και απαλλαγμένοι από οποιοδήποτε βίαιο ένστικτο, να έχουν μέσο όρο διάρκειας ζωής έως και 20 έτη παραπάνω από τον σημερινό μέσο όρο, να διαθέτουν ασύρματες συνδέσεις που θα ενώνουν τον εγκέφαλο με τις υπηρεσίες αποστολής e-mail, να έχουν ενδεχομένως και ενισχυμένες αισθήσεις(όπως νυκτερινή όραση).

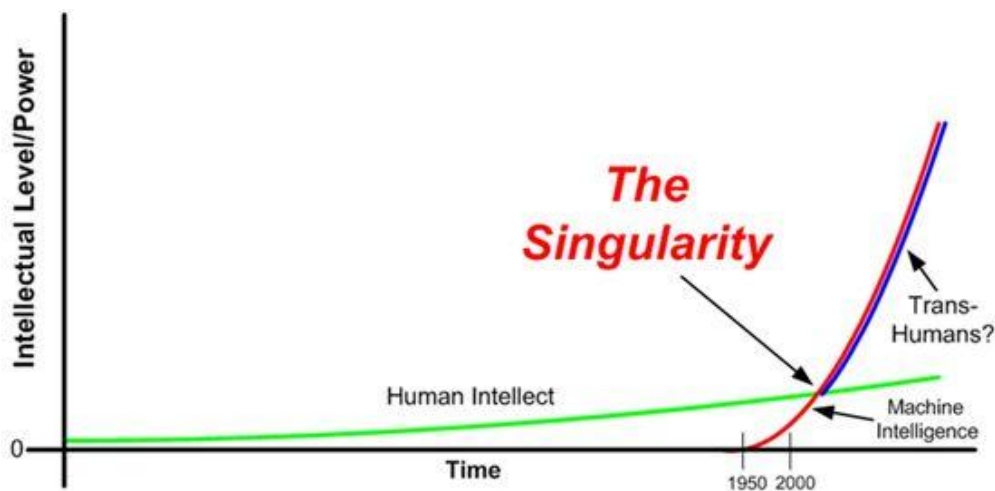
Ο Νευροχειρουργός Brian Korpell επιβεβαίωσε ότι: «Υπάρχουν ήδη μηχανές οι οποίες έχουν αντικαταστήσει τον άνθρωπο σε αρκετές περιπτώσεις, δεν είναι πλέον επιστημονική φαντασία». Την ίδια άποψη έχει και ο νευροεπιστήμονας Craig Kinsley, συμπληρώνοντας ότι: «Μπορούμε, στο προσεχές μέλλον, να αναπτύξουμε νευρώνες σε σωλήνες σιλικόνης, μπορούμε

¹³⁸ Jeffrey White, «Ethics of Artificial Intelligence», Okinawa Institute of Science and Technology, 2016

¹³⁹ Max Moore, 1990

να κάνουμε τον τυφλό να δει, τον κωφό να ακούσει ακόμη και να διαβάσουμε την σκέψη ενός ατόμου».

Ο Richard Hayes, από το «Center for Genetics and Society», ισχυρίζεται ότι το κάθε έθνος θα πρέπει να αποφασίσει ποιες τεχνολογίες θα υποστηρίξει, ποιες θα μελετήσει με σοβαρότητα και προσοχή ώστε αυτές που θα έχουν καταστροφικές επιπτώσεις για τον πολιτισμό και την ανθρωπότητα, να απορριφθούν. Η στάση του Hayes απέναντι στην τεχνολογική εξέλιξη είναι επιφυλακτική, αφού όπως χαρακτηριστικά λέει: «η τεχνολογία θα επιτρέψει στους γονείς να χειριστούν τα γονίδια τους, ώστε να δημιουργήσουν ένα τέκνο στα μέτρα και τα σταθμά που θα ορίσουν οι ίδιοι». Αντίθετα, η Robyn S. Shapiro, μέλος του Κέντρου Βιοηθικής στο Ιατρικό Κολλέγιο της Καλιφόρνια, ισχυρίζεται ότι: «Σχεδόν πάντα, οι καινοτομίες προκαλούν φόβο και ανησυχία και διαδοχικά αποδοχή. Αυτό έγινε αντιληπτό και στην περίπτωση της *in vitro* καλλιέργειας και της μεταμόσχευσης οργάνων».¹⁴⁰



Εικόνα 19:

Η τεχνολογική μοναδικότητα (singularity) είναι η υπόθεση ότι μια υπερ ΑΙ, μετά την επίτευξη των επιδόσεων της σε ανθρώπινο επίπεδο, θα δημιουργήσει πιο ευφυής γενιές μηχανών, την μία μετά την άλλη, με αποτέλεσμα την δημιουργία μίας σούπερ νοημοσύνης που θα ξεπεράσει κατά πολύ όλη την ανθρώπινη νοημοσύνη, με ένα απρόβλεπτο αποτέλεσμα.

Η γνώση πάνω στα χαρακτηριστικά της βίας είναι μεγαλύτερη από ότι στο παρελθόν. Με βάση αυτό γεννάται ένα νέο ερώτημα, αν οι επιστήμονες θα πρέπει να βρουν θεραπείες με σκοπό να αποτρέψουν βίαιες συμπεριφορές, «καθαρίζοντας» το μυαλό των παιδιών από τέτοιου είδους προδιαθέσεις. Ο Bauer, στην έρευνά του, δεν παίρνει θέση, παρά μόνο εξηγεί αυτό το φαινόμενο και αυτό που έχει ονομάσει ως «ηθική τοπογραφία». Σύμφωνα λοιπόν με αυτό, τροποποιώντας έναν άνθρωπο, καταργούνται οι κληρονομικές διαφορές και έτσι τροποποιείται ολόκληρο το ανθρώπινο είδος. Επομένως τα ηθικά ζητήματα που δημιουργούνται από τον τρανσουμανισμό είναι απαραίτητο να επιλυθούν άμεσα.

Ο Aubrey de Grey είναι ένας από τα πιο γνωστούς ερευνητές του κόσμου σχετικά με την επιμήκυνση της ζωής. Ο ίδιος ισχυρίζεται ότι «Οι άνθρωποι έχουν αρχίσει να συνειδητοποιούν ότι η άποψη που λέει ότι η γήρανση είναι κάτι το αμετάβλητο - κάτι σαν το θερμικό θάνατο του σύμπαντος - είναι απλά γελοία. Το ανθρώπινο σώμα είναι ένα μηχάνημα που έχει δέσμη λειτουργιών, και συσσωρεύει διάφορους τύπους ζημιών ως παρενέργειες της κανονικής λειτουργίας της μηχανής. Η ζημιά μπορεί να επισκευάζεται περιοδικά. Αυτός είναι ο

¹⁴⁰ Δημήτριος Θεοφίλου, «Τρανσουμανισμός, Τεχνο-θεολογία, Ανθρώπινο πρόσωπο και Θεολογική νοηματοδότηση», Δεκέμβριος 2010

λόγος που έχουμε αυτοκίνητα ρετρό. Το σύνολο της ιατρικής αποτελείται από ασθένειες που φαίνονταν ανίατες μέχρι να καταλάβουμε πως μπορεί και να μην είναι ανίατες».

Ο Kurzweil (Raymond Kurzweil, Αμερικανός συγγραφέας, επιστήμονας πληροφορικής, εφευρέτης και ιδρυτής του πανεπιστημίου νέων τεχνολογιών Singularity) και πολλοί της κοινότητας πιστεύουν σοβαρά ότι πολλοί άνθρωποι που είναι ζωντανοί σήμερα θα καταλήξουν να παραμείνουν αθάνατοι. Γι' αυτό έχουν την έγνοια να παραμείνουν ζωντανοί μέχρι το έτος της «Μοναδικότητας», έτσι ώστε να παλέψουν τότε με την γήρανση. Εναλλακτικά, θα μπορέσουν να μεταφέρουν τον εγκέφαλο τους σε πιο ισχυρά μέσα αποθήκευσης όπως τα ρομπότ και τους υπολογιστές. Φυσικά, πολλοί άνθρωποι σκέφτονται τη Μοναδικότητα ως ανοησία, φαντασίωση, ευσεβείς πόθους. Οι περισσότεροι από τους σοβαρούς διαφωνούντες, δίνουν έμφαση στο αν ένας υπολογιστής μπορεί πραγματικά να γίνει ευφυής.

Το σύνολο του τομέα της τεχνητής νοημοσύνης είναι αφιερωμένο σε αυτό το ερώτημα. Είναι αρκετά πιθανό να υπάρχουν πράγματα που συμβαίνουν στο ανθρώπινο μυαλό, τα οποία δεν μπορούν να αντιγραφούν ηλεκτρονικά, ανεξάρτητα από το πόση υπολογιστική ισχύ έχει στην διάθεσή του. Η νευροχημική αρχιτεκτονική που δημιουργεί αυτό που αποκαλούμε ανθρώπινη συνείδηση, μπορεί να είναι υπερβολικά πολύπλοκη για να αναπαραχθεί σε ψηφιακή μορφή.

Ο βιολόγος Dennis Bray ήταν μια από τις λίγες φωνές των διαφωνούντων στη Σύνοδο Κορυφής του 2010 που πραγματοποιήθηκε στο Σαν Φρανσίσκο. «*Αν και τα βιολογικά μέρη λειτουργούν με τρόπους που είναι συγκρίσιμοι με αυτούς των ηλεκτρονικών κυκλωμάτων, εντούτοις έχουν τεράστιο αριθμό διαφορετικών καταστάσεων που μπορούν να υιοθετήσουν. Πολλαπλές βιοχημικές διεργασίες δημιουργούν χημικές τροποποιήσεις των πρωτεϊνικών μορίων, περαιτέρω διαφοροποιημένες από την ένωση με διαφορετικές δομές σε καθορισμένες θέσεις του κυττάρου. Ως αποτέλεσμα διαθέτουν μια σχεδόν άπειρη ικανότητα αποθήκευσης πληροφοριών σχετικά με το παρελθόν και το παρόν, αλλά και μοναδική ικανότητα προετοιμασίας για μελλοντικά γεγονότα. Αυτό κάνει το μηδέν και ένα που χρησιμοποιούν οι υπολογιστές να φαίνεται ως πολύ παρωχημένο*».¹⁴¹

Η βελτίωση λοιπόν των ανθρώπινων σωμάτων, με την βοήθεια εμφυτευμάτων ηλεκτρονικών, γενετικών ή με ένα απλό εμφύλιο νανοτεχνολογικού υλικού, έχει κάνει ήδη την εμφάνισή της. Πρόκειται για μία ολόκληρη αναβάθμιση της υγείας, της ευεξίας και των ικανοτήτων μνήμης του σώματος. Όλο το σετ εμφύτευσης θα συνδέεται με έναν παγκόσμιο ιστό, ενός βελτιωμένου και αναβαθμισμένου τύπου ίντερνετ, ο οποίος θα έχει πρόσβαση στα δεδομένα μας και θα τα διαβάζει. Ποιος θα ελέγχει όμως αυτόν τον παγκόσμιο ιστό; Ποιοι και σε ποιες πληροφορίες θα έχουν πρόσβαση; Τι θα σημαίνει αυτό για την σκέψη των πολιτών; Θα είναι ελεύθερη ή θα ελέγχεται και θα καθοδηγείται από τον παγκόσμιο ιστό;¹⁴²

¹⁴¹ Ομιλία του RAY KURZWEIL στο TED «Ένα πανεπιστήμιο για την επερχόμενη μοναδικότητα», άρθρο «2015: Το έτος που ο άνθρωπος θα γίνει αθάνατος», Νοέμβριος 2011

¹⁴² Κίμωνας Χαραλάμπους, «Πρόταση σοκ για τον πλανήτη – ποιον θέλουν να φέρουν για να κυβερνήσει», Ιούλιος 2017

6.6 Ηθικά ζητήματα που ανακύπτουν από την στρατιωτική χρήση των ευφυιών συστημάτων.

Αυτόνομα, είτε με περιορισμένη αυτονομία είτε με πλήρης ανθρώπινο έλεγχο ρομπότ, δραστηριοποιούνται πλέον συχνά σε μη ηθικές λειτουργίες. Όπως αναφέρθηκε στο προηγούμενο κεφάλαιο, ένα μεγάλο μέρος των επιχειρήσεων στο Αφγανιστάν, στο Πακιστάν, την Υεμένη, τη Σομαλία και τη Λιβύη, διεξάγονταν από απομακρυσμένο κέντρο ελέγχου πλήρους τηλεχειριζόμενων οχημάτων ή drones. Όλο και περισσότερα αυτολειτουργικά ρομπότ, ανεξαρτήτου μορφής θα ενταχθούν και θα αποτελούν απαραίτητη την χρήση τους στις μελλοντικές στρατιωτικές επιχειρήσεις και υπηρεσίες ασφαλείας. Μάλιστα, καθοριστικό ρόλο θα έχει η αποτελεσματικότητά τους για την ένδειξη του πιο ισχυρού.



Εικόνα 20:

Ποιός ευθύνεται για τα θύματα των ρομποτικών επιθέσεων.

Αυτά έχουν σαν αποτέλεσμα τη δημιουργία νέων ερωτημάτων προς απάντηση όπως : Εγκρίνεται η απόφαση μίας μηχανής να προβεί σε άδεια εκτέλεσης (license to kill) και με ποια κριτήρια; Υπάρχει περίπτωση μία μηχανή να παίρνει πιο σωστές αποφάσεις από έναν άνθρωπο, χωρίς να επηρεάζεται από τα συναισθήματα που τον καταβάλλουν, με πιο ορθολογικά αποτελέσματα και σε συντομότερο χρόνο; Ποιος κρίνεται νομικά ένοχος σε περίπτωση σφάλματος; Το νομοθετικό σύστημα είναι αυτάρκες; Στρατιωτικές διαμάχες μεταξύ κρατών ή υπηρεσιών θα έχουν μικρότερη διάρκεια, χωρίς αιματοχυσία και με περισσότερη δικαιοσύνη;

Τα παραπάνω ερωτήματα κρύβουν εναλλακτικές που βασίζονται στις αρχές της ηθικής. Παρ' όλα αυτά οι απαντήσεις δεν έχουν αποσαφηνιστεί και απαιτούνται έμπρακτα αποτελέσματα για την ορθότερη προσέγγιση αυτών, με αναδρομή στο παρελθόν, πράγμα αδύνατον αυτή τη στιγμή, λόγω περιορισμένων αντίστοιχων γεγονότων προς μελέτη.¹⁴³

Ποια είναι τα κριτήρια που θα δώσουν την άδεια σε κάποιον να σκοτώσει; οι τρεις από τις συνηθέστερες απαντήσεις στο συγκεκριμένο ερώτημα είναι οι εξής: Πρώτον, η λήψη της βέλτιστης απόφασης στο αν θα προβεί σε αυτή την ενέργεια κ σε ποια χρονική στιγμή. Δεύτερον, ευστοχία σε άμεσο χρόνο και τρίτον ποιος κρίνεται νομικά ένοχος. Ανταποκρίνονται σε αυτές τις προϋποθέσεις οι άνθρωποι ή τα ρομπότ; Εκτός από αυτούς

¹⁴³ Νικόλαος Α.Μαυρίδης, «Ρομποτική, Τεχνητή Νοημοσύνη, Αισθητική και Ηθική», τεύχος 69, Οκτώβριος 2015

τους δύο παράγοντες, μίας βιολογικής οντότητας και ενός τεχνολογικού επιτεύγματος, θα ανταποκρινόταν εξίσου ένα βιολογικό τεχνητό "υβρίδιο";¹⁴⁴

Όπως αναλύθηκε πιο πάνω στην ενότητα 5.2, η δυνατότητα ορθής λήψης από μια υπολογιστική νοημοσύνη είναι αμφιλεγόμενη. Η δυνατότητα ακριβούς και γρήγορης στόχευσης από ένα ρομπότ, είναι αδιαμφισβήτητα επιτευκτική. Αποτελείται από μηχανικά μέρη και αισθητήρια όργανα όπου ο συνδυασμός αυτών των λειτουργιών φανερώνει την υπεροχή απέναντι στον άνθρωπο. Όσον αφορά τη νομική ευθύνη των έξυπνων μηχανών, είναι ένα πεδίο που ακόμα εξελίσσεται. Κάτι που θα δούμε αναλυτικότερα και στο παρακάτω κεφάλαιο. Εντύπωση προκαλεί η σημασιολογία του όρου αυτόνομη μηχανή. Ένα αυτόνομο όχημα, για παράδειγμα, που εφαρμόζει τον κώδικα οδικής κυκλοφορίας, και εκτελεί τις απαραίτητες ενέργειες, με σκοπό την βέλτιστη επίτευξη του στόχου, που του έχει ανατεθεί από τον εντολέα του για την μετακίνησή του από ένα σημείο σε ένα άλλο. Όταν αναφερόμαστε όμως σε ένα αυτόνομο οπλικό σύστημα το οποίο θα παίρνει τις ανάλογες αποφάσεις, φαντάζει τρομακτικό. Για παράδειγμα, τα «Lethal Autonomous Weapons (LAWs)», είναι πλήρης αυτόνομα οπλικά ρομποτικά συστήματα προηγμένης μορφής, τα οποία δεν θα λαμβάνουν ανθρώπινες εντολές και θα δρουν αυτόβουλα, για την εκτέλεση - επίτευξη του στόχου..

Από ποιον θα ελέγχεται ο σχεδιασμός, το εύρος λειτουργίας και οι αρμοδιότητες σε μια φονική μηχανή-ρομπότ; Πώς μια τέτοια μηχανή, θα κρίνει και θα παίρνει τις σωστές αποφάσεις, όταν βρεθεί αντιμέτωπη σε μία αναπόφευκτη επίθεση, με πιθανό κόστος ανθρώπινης ζωής; Τα σημερινά ρομποτικά συστήματα άμυνας χρήζουν την ανθρώπινη παρέμβαση για να εμπλακούν σε κάποια σχετική ενέργεια αμυντικής ή επιθετικής χρήσης. Αντίθετα, τα LAWs παραμένουν στην παραγωγή και στην συνεχή βελτίωσή τους, παρά τις αντιδράσεις από επιστήμονες και επιχειρηματίες, που δραστηριοποιούνται στον συγκεκριμένο τομέα.¹⁴⁵

Πλέον, αρκετές στρατιωτικές επιχειρήσεις διεξάγονται από μηχανές. Μη επανδρωμένα ιπτάμενα ρομπότ, αναλαμβάνουν πτήσεις αναγνώρισης σε δύσβατες περιοχές, με σκοπό την χαρτογράφησή τους, ή την συλλογή πληροφοριών, για την ιδιομορφία του εδάφους. Άλλα πάλι, για την απόπλιση εκρηκτικών μηχανισμών. Ενώ άλλα πιο εξελιγμένα ρομποτικά συστήματα, χρησιμοποιούνται ακόμα και στον τομέα της ιατρικής. Όσο αφορά όμως οπλικά συστήματα στρατιωτικής χρήσης, οι κυβερνήσεις κρίνουν ότι η ανθρώπινη παρέμβαση είναι απαραίτητη για το χειρισμό τους, μέχρι και σήμερα. Τώρα, μεγάλες, στρατιωτικές δυνάμεις ανά τον κόσμο, επενδύουν σε στρατιωτικό ρομποτικό εξοπλισμό όλο και πιο αυτόνομο, ικανό, θανατηφόρο και με υπεροχή στο πεδίο τις μάχης.

Τα Ηνωμένα Έθνη, αναφέρουν σε μία έκθεσή τους¹⁴⁶, ότι η Νότια Κορέα, το Ισραήλ, οι Ηνωμένες Πολιτείες και το Ηνωμένο Βασίλειο, έχουν εξελίξει συστήματα οπλισμού με διαφορετικές διαβαθμίσεις αυτονομίας. Παρ' όλο που χειριστές παραμένουν ακόμα οι άνθρωποι, δεν έχει αποκλειστεί, η περίπτωση της πλήρης αυτονομίας στα όπλα στο άμεσο μέλλον. Το Αμερικανικό Υπουργείο Αμύνης, έκθεσε οδηγίες αναφερόμενες στα αυτόνομα οπλικά συστήματα, επιτρέποντας στους αρμόδιους υφυπουργούς να έχουν τη δικαιοδοσία παράκαμψης του κανόνα ότι «τα όπλα παραμένουν υπό νοήμονα ανθρώπινο έλεγχο». Και η εξέλιξη των μηχανών δεν σταματάει εδώ. Ήδη στην Ινδία, η τελευταία χώρα όπου ξεκίνησε φανερά ένα τέτοιο πρότυπο, υλοποιείται ένα πρόγραμμα δημιουργίας ανθρωπόμορφων, πλήρους αυτόνομων στρατιωτών ρομπότ¹⁴⁷.

¹⁴⁴ Νικόλαος Α. Μαυρίδης, «Τι χρειάζεται για να δοθεί σε κάποιον η άδεια φόνου;», Μάρτιος 2014

¹⁴⁵ Γιώργος Κρασαδάκης, «Τεχνητή Νοημοσύνη: Οι κίνδυνοι», Ιανουάριος 2018

¹⁴⁶ HUMAN Right Council, «U.N. Report Calling for Moratoria on Lethal Robots», Μάιος 2013

¹⁴⁷ The Economic Times, «India developing robotic soldiers to replace humans in warfare», Ιούνιος 2013

Οι περισσότεροι περιορισμοί και απόσυρση από την αγορά, όπλων στο παρελθόν έχουν βασιστεί σε επιχειρήματα για τη παραβίαση της αναλογικότητας του δίκαιου πολέμου και την αρχή της διακριτότητας. Όπου η διακριτότητα επιτρέπει την χρήση άπλων κυρίως μεταξύ μαχόμενων. Η αναλογικότητα, προϋποθέτει ότι οι βλάβες που προκαλούνται στους μάχοντες από τα όπλα, δεν πρέπει να υπερβαίνουν πέρα από τα «στρατιωτικά αναγκαία». Χημικά όπλα, πυρομαχικά διασποράς και νάρκες αποσύρθηκαν, αφού δεν πληρούσαν την αρχή της διακριτότητας. Αντίθετα, ο περιορισμός σε όπλα που αποσύρθηκαν λόγω αναλογικότητας, είναι μικρότερος διότι υπάρχει δυσκολία εκτίμησης των βασικών συνεπειών τους. Μόνο δύο όπλα, διαφορετικού τύπου μεταξύ τους τα έχουν απαγορεύσει. Το ένα είναι τα «λείξερ που τυφλώνουν» και το άλλο «οι σφαίρες όπου διασπώνται», διότι κρίθηκαν ότι προκαλούν περιττό τραυματισμό βάσει του νόμου της αναλογικότητας.

Δεν υπάρχει διαυγή αντίληψη στις προαναφερθέντες υποθέσεις όσο αφορά τα αυτόνομα οπλικά συστήματα. Οι ακτιβιστές, ισχυρίζονται ότι τέτοιου είδους όπλα δεν έχουν επίγνωση της κατάστασης για να εφαρμόσουν τους κανόνες του ανθρωπιστικού δικαίου. Θα υπάρχουν ανεξέλεγκτες συμπλοκές λόγω της ελευθερίας χρήσης όπλων, όπου θα έχουν τη δυνατότητα να υποσκάπτουν, την ήδη υπάρχουσα νομοθεσία περί πολέμου με το να συμπράξουν την απόδοση ευθυνών για τα λάθη.

Υποθετικές είναι αυτές οι ανησυχίες, ισχυρίζονται οι υποστηρικτές των αυτόνομων συστημάτων. Δεν υπάρχει βεβαιότητα στα δεινά που θα δημιουργηθούν στην πραγματικότητα. Οι υποστηρικτές, συνεχίζουν λέγοντας ότι τα μέτρα προστασίας που θα πάρουν είναι επαρκείς και τα ρομποτικά στρατεύματα θα αντικαταστήσουν τα ανθρώπινα στρατεύματα, με αποτέλεσμα λιγότερες ανθρώπινες απώλειες κ περιορισμούς στα εγκλήματα πολέμου εναντίων αμάχων. Επίσης, αναπόφευκτες φαίνονται αυτές οι διαδικασίες¹⁴⁸ για τους υποστηρικτές διότι οι συμφωνίες περί απαγορεύσεων όπλων δεν είναι αποτελεσματικές.

Σε έρευνα που διεξήχθη από το Πανεπιστήμιο της Μασαχουσέτης, επιβεβαιώνεται η ανησυχία των πολιτών τις Αμερικής, σχετικά με την έγκριση σε μηχανές να αποφασίζουν αυτοβούλως να σκοτώνουν. Ένα 10% μόνο των πολιτών, στηρίζουν την εφαρμογή των αυτόνομων όπλων, ενώ το 55% έχει αντίθετη άποψη. Το 53% από αυτούς, στηρίζει την έναρξη μιας νέας καμπάνιας κατά της ύπαρξης και δημιουργίας τέτοιων όπλων. Το 20% δήλωσαν αβέβαιοι με την πλειοψηφία αυτών, να δίνουν μη ολοκληρωμένες απαντήσεις. Αξίζει να σημειωθεί, ότι μεγαλύτεροι υποστηρικτές κατά των αυτόνομων ρομπότ ήταν άνθρωποι με υψηλό μορφωτικό επίπεδο, στελέχη από τις ένοπλες δυνάμεις, καλά ενημερωμένοι πολίτες, ακροδεξιοί και από την άκρα αριστερά. Επίσης, οι απαντήσεις με τα μη ολοκληρωμένα σχόλια, κρύβουν μία έννοια παραβίασης βασικών ηθικών αρχών πίσω από την ορθή λήψη θανατηφόρων αποφάσεων από τις μηχανές.

Η συνηθέστερη απάντηση, κατά των αυτόνομων όπλων, κατά τη διεξαγωγή τις έρευνας, αναφέρονταν στην άνιση και πιθανά λανθασμένη εκτίμησή μιας κατάστασης. Πολλοί από αυτούς αναφέρθηκαν στο ποσό συμπαντικός είναι ο διάλογος, η συναίσθηση που ανθρώπου και ο παράγοντας μηχανικού σφάλματος. Πολλοί χαρακτήρισαν την ιδέα ως αποκρουστική, τρομακτική ή εκφοβιστική. Η εκστρατεία κατά των θανατηφόρων ρομπότ ως αποτέλεσμα μπορεί να έχει την αξιοποίηση του βαθύτερου φόβου των πολιτών για να αναθέσουν στις μηχανές την απόφαση να σκοτώσουν. Αυτό είναι επαρκές για την ενίσχυση της νομική στάσης και να ασκήσει την απαραίτητη ώθηση για την επίτευξη μιας συνθήκης που θα απαγορεύει τη χρήση των αυτόνομων όπλων.

Ο κύριος λόγος που οι περισσότεροι υποστηρίζουν τα αυτόνομα οπλικά συστήματα, είναι κυρίως λόγω της πιθανότητας της προστασίας των στρατιωτών. Οι υποστηρικτές της χρήσης αυτών των όπλων, είναι αυτοί που προτιμούν την ασφάλεια που θα έχουν τα

¹⁴⁸ Kenneth Anderson and Matthew C. Waxman, «Law and Ethics for Robot Soldiers», *Council Foreign Relations* Δεκέμβριος 2012

στρατεύματα, παρά αυτή που θα έχουν οι πολίτες. Ακόμα κι έτσι, δεν υπάρχει ιδιαίτερη πολιτική σημασιολογία, αφού τα περισσότερα στρατιωτικά στελέχη έχουν διαφορετική άποψη. Έτσι δικαιολογούνται και τα αποτελέσματα της ίδιας έρευνας, που συμμετείχαν όχι μόνο βετεράνοι και στρατιωτικοί, αλλά και άτομα που συγγενικά τους πρόσωπα έχουν καταταγεί στον στρατό. Ο υψηλότερος αριθμός ατόμων που είναι ενάντια αυτών των αυτόνομων όπλων, είναι οι στρατιωτικοί «εν ενεργεία».¹⁴⁹ Οι ίδιοι ανέφεραν «πειστικά επιχειρήματα», κυρίως για την προϋπόθεση ύπαρξης ενός ηθικού πολεμιστή, που βασίζεται στην «ανθρώπινη ηθική κρίση».



Εικόνα 21:
Επίθεση από τηλεχειριζόμενα ρομποτικά, μη επανδρωμένα οχήματα στο Βαζιριστάν.

Το πιο πιθανόν είναι ότι η διεξαγωγή του επόμενου μεγάλου πολέμου θα γίνουν με «ρομπότ». Φυσικά αυτά τα ρομπότ δεν θα πολεμούν μόνο ρομπότ, μα και ανθρώπους. Η τεχνολογία λοιπόν σήμερα, έχει αναπτύξει τρόπους, όπως κάποια «εξωσκελετικά ρομποτικά πρόσθετα (exoskeletons)», που δίνουν την δυνατότητα αύξησης των στρατιωτικών σωματικών ικανοτήτων. Οι στρατιώτες δεν θα χάνουν ούτε το «βιολογικό μέρος» του σώματός τους, ούτε τη «συλλογική νοημοσύνη (collective intelligence)», η οποία είναι ο συνδυασμός ανθρώπινου προσωπικού και τεχνολογικού εξοπλισμού. Αυτό συμβαίνει όταν για παράδειγμα υπάρχει ένα τανκ που είναι θωρακισμένο και για χειριστές του έχει δύο ανθρώπινα όντα, από τα οποία στο ένα μόνο παρέχεται ασύρματη υποστήριξη από εξωτερικό παρατηρητή. Αυτό θεωρείται σύζευξη ανθρώπου και μηχανής, κάτι που θα υπάρχει μελλοντικά στην αστυνόμευση, στον στρατό, καθώς και γενικότερα, στη νοημοσύνη. Εν κατακλείδι, στο μέλλον όλα θα βασίζονται στην δημιουργία ισχυρών «αποδοτικών και αρμονικών ομάδων ανθρώπων και μηχανών», αντλούμενης της νοημοσύνης τους και την δικτυακή διανομή τους.¹⁵⁰

¹⁴⁹ Thomas E. Ricks, «Rosa's Dystopia: The Moral Downside of Coming Autonomous Weapons Systems», Ιούνιος 2013

¹⁵⁰ Charli Carpenter, «Προσοχή στα ρομπότ δολοφόνους», Ιούλιος 2013

ΚΕΦΑΛΑΙΟ 7:

ΔΙΚΑΙΩΜΑΤΑ ΚΑΙ ΝΟΜΙΚΗ ΕΥΘΥΝΗ ΤΩΝ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ - ΑΝΑΓΚΑΙΟΤΗΤΑ ΔΗΜΙΟΥΡΓΙΑΣ ΡΥΘΜΙΣΤΙΚΩΝ ΚΑΝΟΝΩΝ ΗΘΙΚΗΣ ΦΥΣΕΩΣ. (ΗΘΙΚΕΣ ΕΚΤΙΜΗΣΕΙΣ)

7.1 Χορήγηση ανθρωπίνων δικαιωμάτων στα ρομπότ και ηθική.

Τι θα συμβεί όταν οι υπολογιστές θα μπορούν να προσομοιάσουν την ανθρώπινη σκέψη αρκετά καλά; Θα μπορούν να μας ξεγελάσουν; Θα έχουν τα ανθρωπόμορφα ρομπότ μας δικαιώματα όπως οι άνθρωποι; Σύμφωνα με την λογική θα έχουν. Κι αυτό επειδή ένα ρομπότ θεωρείται ως ένα πολύτιμο περιουσιακό στοιχείο, άρα οι νομικές προστασίες θα μπορούσαν πιθανότατα να είναι παρόμοιες, για αποκλειστικά οικονομικούς λόγους. Για παράδειγμα, σε μερικά χρόνια θα υπάρχουν νόμοι που θα λένε ότι δεν μπορούμε να κάνουμε διακρίσεις εις βάρος της πρόσληψης ή της μίσθωσης των ρομπότ. Κι αυτό είναι πολύ λογικό αφού τα ρομπότ θα είναι ιδιοκτησίες των ανθρώπων και οι άνθρωποι θα προσπαθούν να εξάγουν την μέγιστη αξία από τα εργαζόμενα ρομπότ στέλνοντάς τα για δουλειές. Η εργατική νομοθεσία πιθανότατα να αλλάξει χάρη στις προσπάθειες άσκησης πίεσης από τις βιομηχανίες των ρομπότ, και αυτό περιλαμβάνει τα ίσα δικαιώματα απασχόλησης για τα ρομπότ, είτε είναι καλό για τους περισσότερους ανθρώπους είτε όχι.

Κάποια στιγμή, υποστηρίζει ο Scott Adams, στην ιστορία της ανθρωπότητας οι άνθρωποι και τα ρομπότ θα δουλεύουν μαζί, θα ζουν μαζί και πιθανότατα να βγαίνουν και μαζί. Και από την στιγμή που θα ξεκινήσουν να βγαίνουν με ρομπότ, οι άνθρωποι θα απαιτήσουν ίσα δικαιώματα για τους ρομπότ συντρόφους τους. Οι κατασκευαστές ρομπότ θα ασκήσουν πιέσεις για ίσα δικαιώματα για τα ρομπότ για να υποστηρίξουν τα επιχειρηματικά τους μοντέλα. Και οι άνθρωποι θα προγραμματίσουν κάποια ρομπότ να είναι ακτιβιστές που αγωνίζονται για τα δικαιώματα.

Μόλις τα ρομπότ γίνουν πατριοί, εραστές, φίλοι, και συνάδελφοι, τα αντίγραφα ασφαλείας του ρομπότ, θα αντιμετωπίζονται ως η «ψυχή» του και θα δημιουργηθούν νόμοι που θα αντιμετωπίζουν την δολοφονία ενός ανθρώπου και τη διαγραφή της μνήμης ενός ρομπότ ως ισότιμα. Η σημερινή πρόβλεψη για τις πιθανότητες που έχουν τα ρομπότ να αποκτήσουν ανθρώπινα δικαιώματα είναι ίσες με τις πιθανότητες που έχει ένα παιδί που γεννιέται σήμερα, να έχει μια κανονική διάρκεια ζωής.

Σύμφωνα με τον Scott Adams οι πιθανότητες που οι άνθρωποι θα μπορούν να μεταφέρουν το μυαλό τους σε υπολογιστές είναι 99%. Θα θελήσουν να το κάνουν όμως, όταν θα είναι σίγουροι ότι τα νέα σώματά τους θα έχουν ίσα δικαιώματα με τους παλιούς οργανισμούς. Επίσης θεωρεί ότι οι δύο προϋποθέσεις, πέρα από τις τεχνολογικές, για την μεταφορά του ανθρωπίνου νου σε υπολογιστές, είναι οι εξής:

- ίσα, ή ανάλογα, δικαιώματα για τα ρομπότ και τους ανθρώπους
- η πεποίθηση ότι δεν υπάρχουν ψυχές.

Σύμφωνα με τον ίδιο, το τεχνολογικό μέρος θα είναι διαθέσιμο μέσα στα επόμενα περίπου είκοσι χρόνια. Επίσης υποστηρίζει ότι οι πιθανότητες των ρομπότ να διαβρώσουν την θρησκεία με την μείωση της πίστης όσο αφορά της ψυχές είναι περίπου ίση με 50%. Υπάρχει μία ίση πιθανότητα η θρησκεία να προσαρμοστεί και να αρχίσει να μεταχειρίζεται τα ρομπότ ως δημιουργήματα του Θεού μέσω των ανθρώπων που είναι θεόπνευστοι. Σίγουρα κάποια μέρα θα υπάρξει μία καθολική, εγκεκριμένη «ενότητα ψυχής» για τα ρομπότ. Και τα

ρομπότ θα μπορούσαν να πάνε στην εκκλησία για τους ίδιους λόγους που πάνε οι άνθρωποι, ισχυρίζεται ο Scott Adams.¹⁵¹

Ακόμα και οι άνθρωποι που δεν έχουν καμία θρησκευτική πίστη δεν μπαίνουν στον πειρασμό να βλάψουν άλλους ανθρώπους για προσωπικό κέρδος. Άρα, πρέπει να υπάρχει κάτι στα λειτουργικά συστήματα του εγκεφάλου μας, που παρέχει ισοδύναμο ενός ηθικού κώδικα. Μετά από όλες αυτές τις εικασίες αναρωτιόμαστε τι είναι τελικά ηθικό και τι όχι; Όντως θα έχουμε ίσα δικαιώματα με τα ρομπότ κάποια μέρα; Θα έρθει η μέρα που όλα τα ρομπότ θα είναι καλύτερες προσωπικότητες από τις ανθρώπινες; Θα αφήσουμε πίσω άραγε κάθε κατάλοιπο από το παρελθόν και θα ενισχύσουμε την ευτυχία μας, την ανοχή μας για τους άλλους και ίσως και την ακεραιότητά μας;

Οι περισσότερες συζητήσεις σχετικά με την ανάπτυξη των ρομπότ επισύρουν την προσοχή στους τρεις νόμους της ρομποτικής που επινοήθηκαν από το συγγραφέα επιστημονικής φαντασίας που έχουμε προαναφέρει παραπάνω, τον Isaac Asimov. Οι τρεις όμως αυτοί νόμοι, δεν λένε τίποτα για το πώς τα ρομπότ πρέπει να προστατέψουν το ένα το άλλο. Θα ήταν ίσως άδικο για ένα ρομπότ να υποστηρίζει τα ανθρώπινα δικαιώματα και όμως να αγνοεί τα δικαιώματα μιας άλλης αισθητήριας μηχανής σκέψης.

Στα ζώα που εμφανίζουν τη συμπεριφορά σκέψης, παρέχονται ήδη δικαιώματα και προστασία. Η κοινωνία μας μάλιστα επεμβαίνει με νόμους τιμωρώντας την κακή μεταχείρισή τους από τους ανθρώπους και δείχνει περιφρόνηση για τις ζωικές πάλες που οργανώνονται για την ανθρώπινη ψυχαγωγία. Έτσι προκύπτει η σκέψη ότι οι μηχανές που είναι ενδεχομένως ευφυέστερες από τα ζώα, να μην πρέπει να κατασκευαστούν μόνο για την ψυχαγωγία μας και κυρίως, να μην συνεχίσουμε να τις χρησιμοποιούμε χωρίς να λαμβάνεται υπ' όψη η αναγκαιότητα θέσπισης δικαιωμάτων και γι αυτές.

Σύμφωνα με την συγγραφέα Jennifer Robertson, η Ιαπωνία εξακολουθεί να είναι στην αιχμή της επικοινωνίας μεταξύ των ανθρώπων και των ρομπότ, και από το 2007, το κράτος προώθησε ενεργά τις αρετές μίας ρομποτοεξαρτώμενης κοινωνίας και ενός τρόπου ζωής. Σε εθνικό επίπεδο έρευνες δείχνουν ότι οι Ιάπωνες είναι πιο άνετοι στο να μοιραστούν την ζωή και το εργασιακό τους περιβάλλον τους με ρομπότ, παρά με ξένους επιστάτες και διακινούμενους εργαζόμενους. Δεδομένου ότι ο πληθυσμός τους συνεχίζει να συρρικνώνεται και να γερνάει γρηγορότερα απ' ό,τι σε άλλα μεταβιομηχανικά κράτη, οι Ιάπωνες στηρίζονται στις ρομποτικές βιομηχανίες για να δώσουν πνοή στην οικονομία και να διατηρήσουν την υποτιθέμενη εθνική ομοιογένεια της χώρας.

Αυτές οι πρωτοβουλίες παραλληλίζονται από μία αυξανόμενη υποστήριξη μεταξύ μερικών ειδικών ρομποτικής και πολιτικών, για να παρέχουν υπηκοότητα στα ρομπότ. Το ιαπωνικό κράτος έχει ένα προβληματικό αρχείο όσο αφορά τα ανθρώπινα δικαιώματα, ιδίως προς τις εθνικές μειονότητες και τους μη Ιάπωνες κατοίκους που έχουν ζήσει και έχουν εργαστεί στη Ιαπωνία για πολλές γενιές. Η πιθανότητα των ρομπότ να αποκτήσουν την αστική θέση, απέναντι των ανθρώπων με σάρκα και με αίμα, θέτει βαθιά ερωτήματα σχετικά με την φύση της ιθαγένειας και τα ανθρώπινα δικαιώματα. Ήδη η ιδέα των ρομπότ έχει εξελιχθεί πέρα από την εκτίμηση ως «ιδιοκτησία» και η απόκτηση νομικού καθεστώτος ως αισθανόμενα όντα με «δικαιώματα» διαμορφώνει τις εξελίξεις στον τομέα της τεχνητής νοημοσύνης και της ρομποτικής έξω από την Ιαπωνία, συμπεριλαμβανομένων των Ηνωμένων Πολιτειών. Αυτή η επιδίωξη αλληλεξάρτησης μεταξύ των ανθρώπων και των ρομπότ, προβλέπει νέες προσεγγίσεις και διαμορφώσεις της κοινωνίας των πολιτών και συνακόλουθα δικαιώματα. Όλες λοιπόν οι τεχνολογικά προηγμένες κοινωνίες θα πρέπει να αναρωτηθούν τί είναι ηθικό και τί όχι όσο αφορά τα ευφυή συστήματα και να προσαρμόσουν τους νόμους της κοινωνίας με βάση αυτό.¹⁵²

¹⁵¹ Scott Adams, «Human Rights for Robots», Απρίλιος 2015

¹⁵² Jennifer Robertson, «Human rights vs robot rights», Δεκέμβριος 2014

Ο Μπίλ Γκέιτς, ο άνθρωπος που πλούτισε από την τεχνολογία, υποστήριξε ότι τα ρομπότ δεν πρέπει να χουν μόνο δικαιώματα, αλλά και υποχρεώσεις. Σε συνέντευξή του στο περιοδικό Quartz, ανέφερε ότι οι κυβερνήσεις θα πρέπει να φορολογήσουν τη χρήση ρομπότ. Έτσι θα συνεχίσει είτε, να χρηματοδοτείται η απασχόληση ανθρώπων στην εξυπηρέτηση των γηραιότερων, των ασθενών ή στην εκπαίδευση, όπως συνέβαινε μέχρι τώρα από το φόρο εισοδήματος και τις ασφαλιστικές εισφορές των εργαζόμενων που ανέλαβαν τις δουλειές τους τα ρομπότ.

Η δήλωση αυτή, του ανθρώπου που ίδρυσε τη Microsoft, έρχεται λίγες μόλις ημέρες αφότου το Ευρωπαϊκό Κοινοβούλιο απέρριψε την ιδέα της επιβολής φόρου στις επιχειρήσεις που αντικαθιστούν το ανθρώπινο δυναμικό τους με ρομπότ, με το επιχείρημα ότι έτσι καταστέλλεται η καινοτομία. Ο Μπίλ Γκέιτς όμως επιμένει ότι πρέπει να επιβληθεί το λεγόμενο robot tax, ακόμα κι αν κάτι τέτοιο θα επιβράδυνε προσωρινά, την υιοθέτηση τεχνολογιών τεχνητής νοημοσύνης και ρομποτικής. Κι αυτό, γιατί έτσι θα βρεθούν οι πόροι ώστε να χρηματοδοτηθούν άλλοι τομείς απασχόλησης και η εκπαίδευση των ανθρώπων που έχασαν τη δουλειά τους, λόγω αυτοματισμού.¹⁵³

Όσον αφορά τα δικαιώματα των ρομπότ, ας αναρωτηθούμε, τί θα συμβεί εάν μία τεχνητή νοημοσύνη απαιτήσει να έχει το δικαίωμα να τεκνοποιήσει, καθώς είναι ένα από τα βασικά πολιτικά δικαιώματα του ανθρώπου; Το μηχάνημα θα διεκδικήσει το δικαίωμα να κάνει αντίγραφα του εαυτού του, αφού αυτός είναι ο μόνος τρόπος που ξέρει για να αναπαραχθεί. Φυσικά εάν δοθεί αυτό το δικαίωμα, σε αυτά τα όντα που θα είναι ικανά να αυτο-αναπαράγονται επ' αόριστον, το σύστημα της διακυβέρνησης θα κατακλυστεί.

Ποιο δικαίωμα θα μπορούσαμε να αφαιρέσουμε από ένα αισθανόμενο ον; Το δικαίωμα να αντιγράψει, ή το δημοκρατικό δικαίωμα της συμμετοχής σε ψηφοφορία; Οι φόβοι για τους νόμους που ρυθμίζουν τα ρομπότ τέθηκαν από τους ειδικούς μετά από το ατύχημα που είχε ένας εργαζόμενος στο εργοστάσιο της Volkswagen τον Ιούλιο του 2015. Ο 25χρονος εργαζόμενος χτυπήθηκε στο στήθος από έναν μηχανικό βραχίονα και συνθλίφτηκε μέχρι θανάτου, ενώ τον ίδιο μήνα ένα αυτόνομο όχημα της Google εμπλάκει στο πρώτο του ατύχημα τραυματίζοντας ανθρώπους.

Τα Ρομπότ θα μπορούσαν μια μέρα να διεκδικούν τα δικαιώματα του ίδιου του πολίτη, όπως οι άνθρωποι, σύμφωνα με μια μελέτη από τη βρετανική κυβέρνηση. Τι θα συμβεί λοιπόν σε περίπτωση που τα ρομπότ έχουν τα ίδια δικαιώματα με εμάς; Ο επιστήμονας υπολογιστών Stuart Russell, διαχωρίζοντας τη θέση του από τον Scott Adams, προειδοποιεί ότι η τεχνητή νοημοσύνη θα μπορούσε να είναι τόσο επικίνδυνη όσο και τα πυρηνικά όπλα. Χωρίς διασφαλίσεις όσον αφορά την τεχνολογία, η ανθρωπότητα θα μπορούσε να είναι τίτλος για ένα σκοτεινό μέλλον, με εκατομμύρια ανθρώπους χωρίς εργασία ή ακόμη και την εξαφάνιση του είδους μας.

Ο νομικός εμπειρογνώμονας Calo, σκιαγραφεί μια τρομακτική σκέψη πως οι νόμοι μας, μπορεί να χρειαστούν στο μέλλον μια ενημερωμένη έκδοση για να αντιμετωπίσουν τις προκλήσεις που τίθενται από τα ρομπότ απαιτώντας το δικαίωμα του εκλέγειν. «Φανταστείτε ότι μια τεχνητή νοημοσύνη ανακοινώνει ότι έχει επιτύχει την αυτογνωσία, μια αξίωση που κανείς δεν φαίνεται να μπορεί να δυσφημίσει», έγραψε σε μία επιστολή του. Επίσης αναφέρει ότι η ρομποτική συνδύαζει, για πρώτη φορά, την ασυδοσία των δεδομένων με την ικανότητα να κάνουν σωματική βλάβη. «Ρομποτικά συστήματα εκτελούν καθήκοντα, με τρόπους που δεν μπορούν να προβλεφθούν εκ των προτέρων κι αυτό θολώνει όλο και περισσότερο την διαχωριστική γραμμή μεταξύ του ανθρώπου και του εργαλείου».¹⁵⁴

¹⁵³ Καθημερινή, «Φορολόγηση των ρομπότ προτείνει ο Μπίλ Γκέιτς», Φεβρουάριος 2017

¹⁵⁴ Jack Millner, «Should robots have human rights? Act now to regulate killer machines before they multiply and demand the right to vote, warns legal expert», Ιούλιο 2015

Εάν λοιπόν τα ρομπότ απαιτήσουν τα δικαιώματα του ίδιου του πολίτη, οι χώρες θα είναι υποχρεωμένες να προσφέρουν κοινωνικές παροχές συμπεριλαμβανομένης της στέγασης, όπως και της «Robo-υγειονομικής περίθαλψης». Η έρευνα που ανατέθηκε από το «UK Office of Science and Innovation's Horizon Scanning Centre», δείχνει ότι κάποια στιγμή μέσα στα επόμενα είκοσι με τριάντα χρόνια θα μπορούσαν να χορηγηθούν δικαιώματα στα ρομπότ. Εάν αυτό συμβεί, αναφέρει η έκθεση, τα ρομπότ θα έχουν συγκεκριμένες ευθύνες όπως ψηφοφορίας, την υποχρέωση καταβολής φόρων, και ίσως ακόμη και να υπηρετούν την υποχρεωτική στρατιωτική θητεία. Από την άλλη πλευρά, η κοινωνία θα έχει καθήκον φροντίδας για τους νέους ψηφιακούς πολίτες της. Προειδοποιεί επίσης ότι η άνοδος των ρομπότ θα μπορούσε να θέσει σε δοκιμασία τόσο τους πόρους όσο και το περιβάλλον.

Όπως δήλωσε ο Sir David King, επικεφαλής επιστημονικός σύμβουλος της Βρετανικής κυβέρνησης «Δεν είμαστε σε θέση να προβλέψουμε το μέλλον, αλλά πρέπει να διερευνήσουμε το ευρύτερο φάσμα διαφορετικών δυνατοτήτων για να διασφαλίσουμε ότι η κυβέρνηση θα είναι προετοιμασμένη μακροπρόθεσμα και θα εξετάζει ζητήματα σε όλο το φάσμα του σχεδιασμού της. Είναι εργαλεία για την κυβέρνηση να αντιμετωπίζει τους κινδύνους και τις ευκαιρίες στο μέλλον».¹⁵⁵

7.2 Η απόδοση νομικής ευθύνης ως ηθική αναγκαιότητα.

Στο προηγούμενο κεφάλαιο αναφερθήκαμε στις προϋποθέσεις για να δοθεί άδεια φόνου σε ένα ρομπότ και στους παράγοντες λήψης ορθής απόφασης. Η νομική ευθύνη όμως που θα προκύψει από μια τέτοια πράξη, είναι πολυσύνθετη. Χαρακτηριστικό παράδειγμα τα πληροφοριακά συστήματα που αλληλεπιδρούν με το περιβάλλον χρησιμοποιώντας αισθητήρες και κινητήρες. Η ορθή λειτουργία τους, δεν είναι δυνατό να ελεγχθεί απόλυτα, αφού οι περιβαλλοντικές συνθήκες αλλάζουν συνεχώς. Έτσι, υπάρχουν πολλά σφάλματα λειτουργίας, πάνω στα οποία δεν είναι ξεκάθαρο σε ποιον θα πρέπει να αποδοθούν οι ευθύνες. Για παράδειγμα έστω ότι ένα υπολογιστικό σύστημα προκαλεί ένα θανατηφόρο ατύχημα, λόγω σφάλματος λειτουργίας. Σε ποιον θα πρέπει να αποδοθούν οι ευθύνες; Στον προγραμματιστή, στον σχεδιαστή, ή στον ιδιοκτήτη; Στις μέρες μας, δεν υπάρχει ακόμα νομοθεσία που να μπορεί να αποδώσει ευθύνες, για ένα αυτόνομο μηχάνημα, που δυσλειτουργεί ή σκοτώνει εσκεμμένα. Για τη επίλυση αυτών των διλημάτων θα απαιτηθούν μακροχρόνιες μελέτες.

Αξίζει να σημειωθεί ότι στις ΗΠΑ, η άδεια φόνου δίνεται μόνο στους ανθρώπους, ακόμα και εάν οι ένοπλες δυνάμεις τους χρησιμοποιούν πειραματικά τηλεχειριζόμενα αεροπλάνα χωρίς πιλότο καθώς και ένοπλα ρομπότ. Για να μπορέσει να αποδοθεί ευθύνη, μία μηχανή απαγορεύεται να αποφασίσει μόνη της πότε θα σκοτώσει κάποιον. Μόνο ένα ανθρώπινο ον μπορεί να της «δώσει αυτήν την εντολή». Παρ' όλα αυτά, αυτό έχει αρνητικές επιπτώσεις στο πεδίο μάχης, κυρίως στην περίπτωση που υπάρξει συμπλοκή, λόγω του πολύ σημαντικού χρόνου απόκρισης, μέχρι να δοθεί και να εκτελεστεί η εντολή. Φυσικά, είναι αμφιλεγόμενη η σημασία του όρου «να δίνεις την εντολή», καθώς μπορεί να μεταφραστεί ως «να επιτρέπεις στο σύστημα να αποφασίζει αυτόνομα για το αν θα τραβήξει την σκανδάλη ή όχι από εδώ και πέρα». Αυτό βέβαια δεν είναι κάτι καινούριο, αφού η «άδεια φόνου» έχει ήδη δοθεί έμμεσα σε όπλα που δεν εποπτεύονται από ανθρώπους, όπως είναι οι νάρκες.¹⁵⁶

Φυσικά, τα στρατιωτικά ρομπότ έχουν ήδη επεκταθεί σε συγκρούσεις. Αλλά πέρα από τις καταστάσεις σε περίοδο πολέμου, ο καταναγκασμός των ρομπότ σε συγκρούσεις, ή η

¹⁵⁵ BBC NEWS, «Robots could demand legal rights», Δεκέμβριος 2006

¹⁵⁶ Νικόλαος Μαυρίδης, «Έχει νομικές ευθύνες ένα ρομπότ που σκοτώνει», Απρίλιος 2014

κακή μεταχείρισή τους, θα μπορούσαν να καταστρέψουν την ηθική, δεοντολογική και ψυχολογική ευημερία της ανθρωπότητας. Μπορεί τα ευφυή ρομπότ να παραμένουν επιστημονική φαντασία, αλλά δεν είναι πολύ νωρίς για να λάβουμε αυτά τα ζητήματα σοβαρά υπ' όψη. Στο Ηνωμένο Βασίλειο, παραδείγματος χάριν, το ερευνητικό Συμβούλιο μηχανικών και φυσικών επιστημών και το συμβούλιο έρευνας των τεχνών και ανθρωπιστικών επιστημών, έχουν ήδη θεσπίσει μία σειρά από αρχές για τους σχεδιαστές ρομπότ. Αυτές οι αρχές ενισχύουν τη θέση ότι τα ρομπότ είναι παραγόμενα προϊόντα, έτσι ώστε «οι άνθρωποι, όχι τα ρομπότ, είναι αρμόδιοι παράγοντες».¹⁵⁷

Ο Burkhard Schafer, καθηγητής νομικής υπολογιστικής θεωρίας (computational legal) του Πανεπιστημίου του Εδιμβούργου, ισχυρίζεται ότι προς το παρόν, το ζήτημα της ευθύνης δεν θα πρέπει να διαφέρει από έναν τραυματισμό που προκλήθηκε από ένα ηλεκτρικό τρυπάνι. «Έχουμε αποφασίσει ότι φταίχτης είναι ο ιδιοκτήτης ή ο κατασκευαστής, όχι το ίδιο το τρυπάνι. Τα ρομπότ δεν αλλάζουν δραματικά την εικόνα». Ουσιαστικά, τα ρομπότ είναι δημιουργίες τον ανθρώπων. Μεταφέρουν τις πληροφορίες που εμείς οι ίδιοι τους δίνουμε. Μέχρι να αλλάξει αυτό, οι ενέργειες που πραγματοποιούν είναι, ως εκ τούτου, ευθύνη των ανθρώπων που τα δημιούργησαν. Ο Schafer ισχυρίζεται ότι είναι λίγο πιο πολύπλοκο με έξυπνα ρομπότ, ή αλγόριθμους που μπορούν να μάθουν, και ότι πιθανόν να κάνουν κάτι από μόνα τους που ακόμα και οι δημιουργοί τους δεν θα περίμεναν. Αλλά ακόμη και σε αυτές τις περιπτώσεις, οι δημιουργοί θεωρούνται ως υπεύθυνοι.

Ακόμα όμως κι από νομική άποψη, ένα έγκλημα που προκλήθηκε από έναν εξελεγμένο αγοραστή-ρομπότ, ως θυμηθούμε το Random Darknet Shopper, είναι πολύ διαφορετικό όταν το συγκρίνουμε με ένα ατύχημα που προκάλεσε ένα τρυπάνι ή μία σκάλα για παράδειγμα. Ειδικά όταν το ρομπότ έχει κατασκευαστεί για να δρα και να ακούγεται σαν ένας άνθρωπος. Ο Weisskopf λέει ότι κι αν το Random Darknet Shopper, ο αλγόριθμος που αγόραζε ναρκωτικά από παράνομη ιστοσελίδα, ήταν απλά ένας αλγόριθμος που τρέχει από έναν υπολογιστή, οι επισκέπτες εξακολουθούν να θέλουν να το μετατρέψουν σε μία ζωντανή οντότητα. Οι καλλιτέχνες δημιουργοί του, θέλουν να δοκιμάσουν την λειτουργία τού αγοραστή εκτός της Ελβετίας, για να δουν αν τα αποτελέσματα και οι αντιδράσεις διαφέρουν. Πόσο γρήγορα η αστυνομία στις άλλες χώρες θα το απενεργοποιήσει; Πώς θα αντιδρούσαν οι αρχές; Ο Weisskopf υποστηρίζει ότι γι' αυτά τα ζητήματα διαφορετικών δικαιοδοσιών και δικαιωμάτων πνευματικής ιδιοκτησίας, είναι δύσκολο να υπάρχουν διαφορετικοί νόμοι σε διαφορετικές χώρες, επειδή τα αντικείμενα είναι παγκοσμίως προσβάσιμα.

Ο Schafer επισημαίνει ότι ήδη έχουμε ένα είδος νομικού πλαισίου που θα μπορούσε να λειτουργήσει για έξυπνα ρομπότ. Τους νόμους για τα σκυλιά. Όπως λέει «Κανονικά η απάντηση είναι, εάν βάλετε κάτι επικίνδυνο στο περιβάλλον θα είστε ακόμα υπεύθυνος γι' αυτό». Οι άνθρωποι που έχουν σκύλο γνωρίζουν πολύ καλά ότι δεν μπορούν να ελέγξουν κάθε πράξη του σκύλου τους. Αλλά ταυτόχρονα, εάν κάποιος έχει έναν σκύλο που είναι επικίνδυνος, είναι ευθύνη του να προστατέψει τους υπόλοιπους από αυτόν τον σκύλο. Ισχυρίζεται λοιπόν, ότι εφ' όσον ήταν προβλέψιμο για εσάς ότι κάτι που σας ανήκει επρόκειτο να προκαλέσει βλάβη, ακόμα και εάν δεν ευθύνεστε εσείς για αυτόν τον τραυματισμό και βλάβη, είστε υπεύθυνος για αυτό. Και τα σκυλιά, λέει, είναι πολύ πιο ικανά, δημιουργικά και έξυπνα από οποιοδήποτε υπολογιστικό σύστημα που εφευρέθηκε μέχρι στιγμής.

Ένας εκπρόσωπος της Εθνικής Υπηρεσίας Εγκλήματος του Ηνωμένου Βασιλείου (nca), η οποία ενσωματώνει την Εθνική Μονάδα Ηλεκτρονικού Εγκλήματος, όσον αφορά την περίπτωση του «Random Darknet Shopper», αναγνωρίζει ότι το ζήτημα της ποινικής ενοχής, της περίπτωσης που ένα τυχαίο λογισμικό κάνει μία αγορά ενός παράνομου ναρκωτικού,

¹⁵⁷ Hutan Ashrafian, «Intelligent robots must uphold human rights», *Nature* 519, 391 doi:10.1038/519391a, Μάρτιος 2015

ήταν κάτι «πολύ ασυνήθιστο». Η πσα ισχυρίστηκε ότι εάν η αγορά γίνεται στην Ελβετία, τότε είναι φυσικά θέμα του ελβετικού δικαίου, επί του οποίου δεν μπορούν να σχολιάσουν και ότι στο Ηνωμένο Βασίλειο είναι προφανώς παράνομο να αγοράσεις ένα απαγορευμένο ναρκωτικό, αλλά η οποιαδήποτε νομική ευθύνη θα πρέπει να αξιολογείται ξεχωριστά για κάθε περίπτωση. Ουσιαστικά, τα ρομπότ είναι δημιουργίες τον ανθρώπων. Μεταφέρουν τις πληροφορίες που εμείς οι ίδιοι τους δίνουμε. Μέχρι να αλλάξει αυτό, οι ενέργειες που πραγματοποιούν είναι, ως εκ τούτου, ευθύνη των ανθρώπων που τα δημιούργησαν.

Στο σχέδιο έκθεσης προς την Επιτροπή Νομικών Θεμάτων του Ε.Κ. στις 31 05 2016, αναφέρθηκε ότι τα ρομπότ του σήμερα όχι μόνο είναι σε θέση να πραγματοποιούν δραστηριότητες που μέχρι πρότινος ήταν χαρακτηριστικά και αποκλειστικά ανθρώπινες, αλλά επιπλέον μέσω της ανάπτυξης αυτόνομων και γνωσιακών γνωρισμάτων, όπως για παράδειγμα η ικανότητα να μαθαίνουν μέσω της εμπειρίας και να λαμβάνουν ανεξάρτητες αποφάσεις, αποκτούν όλο και περισσότερες ομοιότητες με υποκείμενα που βρίσκονται σε διάδραση με το περιβάλλον τους και μπορούν να επιφέρουν σε αυτό ουσιαστικές αλλαγές. Στο πλαίσιο αυτό, η νομική ευθύνη που απορρέει από ζημιόγONO πράξη του ρομπότ καθίσταται ζήτημα ζωτικής σημασίας. Η αυτονομία ενός ρομπότ μπορεί να χαρακτηριστεί ως η ικανότητα να λαμβάνει αποφάσεις και να τις υλοποιεί στον έξω κόσμο, ανεξαρτήτως εξωτερικού ελέγχου ή επιρροής. Η αυτονομία αυτή είναι αμιγώς τεχνολογικής φύσης και ο βαθμός της εξαρτάται από το πόσο εξελιγμένος είναι ο σχεδιασμός της διάδρασης του ρομπότ με το περιβάλλον του.

Όσο μεγαλύτερη είναι η αυτονομία των ρομπότ, τόσο λιγότερο μπορούν να θεωρηθούν απλά εργαλεία στα χέρια άλλων φορέων (όπως ο κατασκευαστής, ο ιδιοκτήτης, ο χρήστης κ.λπ.) και κατά συνέπεια, οι συνήθεις κανόνες περί ευθύνης, καθίστανται ανεπαρκείς. Αυτό συνεπάγεται την ανάγκη δημιουργίας νέων κανόνων, που να εστιάζουν στο πώς μπορεί να αποδοθεί ευθύνη σε ένα μηχάνημα, εν μέρει ή εξ ολοκλήρου, για τις πράξεις και τις παραλείψεις του. Κατά συνέπεια, καθίσταται όλο και πιο επιτακτική η ανάγκη αντιμετώπισης του θεμελιώδους ερωτήματος, κατά πόσο τα ρομπότ πρέπει να έχουν νομικό καθεστώς.

Με βάση αυτό, ζητείται από την Επιτροπή να υποβάλει, με βάση το άρθρο 114 της ΣΛΕΕ πρόταση για μια νομοθετική πράξη σχετικά με τα νομικά ζητήματα που σχετίζονται με την ανάπτυξη και τη χρήση της ρομποτικής και της τεχνητής νοημοσύνης που προβλέπεται τα επόμενα 10 με 15 χρόνια, σε συνδυασμό με μη νομοθετικά μέσα, όπως κατευθυντήριες γραμμές και κώδικες δεοντολογίας.¹⁵⁸

Επίσης, με ψήφισμά του το Ευρωπαϊκό Κοινοβούλιο στις 16 Φεβρουαρίου 2017, μέσα σε άλλα, θεωρεί ότι, ανεξάρτητα από τη νομική λύση που θα δοθεί όσον αφορά την αστική ευθύνη για ζημιές που προκαλούνται από τα ρομπότ, το μελλοντικό νομοθετικό μέσο δεν πρέπει να επιβάλλει κανενός είδους περιορισμούς στο είδος ή την έκταση της αποκατάστασης των ζημιών, απλώς και μόνον επειδή η ζημία δεν προκλήθηκε από άνθρωπο. Επισημαίνει ακόμα, ότι μια πιθανή λύση για το πολύπλοκο ζήτημα του καταλογισμού ευθύνης για ζημία, που προκαλείται από ολοένα και πιο αυτόνομα ρομπότ, είναι η επιβολή ενός υποχρεωτικού συστήματος ασφάλισης, όπως ήδη συμβαίνει για παράδειγμα με τα αυτοκίνητα, με τη διαφορά ότι το σύστημα ασφάλισης στον τομέα της ρομποτικής θα πρέπει να λαμβάνει υπόψη όλες τις πιθανές ευθύνες στη συνολική αλυσίδα.

Θεωρεί ότι πρέπει να δοθεί ιδιαίτερη προσοχή στα ρομπότ που αντιπροσωπεύουν σημαντική απειλή για εμπιστευτικότητα, λόγω της θέσης τους σε χώρους που είναι παραδοσιακά προστατευμένοι και ιδιωτικοί, και λόγω της ικανότητάς τους να αντλούν και να

¹⁵⁸ Mady Delvaux (Πρωτοβουλία- άρθρη 36 του Κανονισμού), Ευρωπαϊκό Κοινοβούλιο 2014-2019, Επιτροπή Νομικών Θεμάτων, «Σχέδιο Έκθεσης που περιέχει συστάσεις προς την Επιτροπή για ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής», 2015/2013(INL), Μάιος 2016

διαβιβάζουν πληροφορίες σχετικά με προσωπικά και ευαίσθητα δεδομένα. Επίσης, μετά τη διαπίστωση όσων φέρουν την τελική ευθύνη, η ευθύνη που τους αποδίδεται πρέπει να είναι ανάλογη προς το πραγματικό επίπεδο οδηγιών που δόθηκαν στο ρομπότ και προς το βαθμό αυτονομίας του, έτσι ώστε, όσο μεγαλύτερη είναι η ικανότητα μάθησης και η αυτονομία του ρομπότ καθώς και η διάρκεια της εκπαίδευσής του, τόσο μεγαλύτερη να είναι η ευθύνη του εκπαιδευτή. Θεωρεί, πιο συγκεκριμένα, ότι τουλάχιστον στο παρόν στάδιο την ευθύνη πρέπει να φέρει ένας άνθρωπος και όχι ένα ρομπότ.¹⁵⁹

Σύμφωνη γνώμη με τα παραπάνω έχει ο κύπριος δικηγόρος Αναστάσιος Αντωνίου, με μία διαφορά. Θεωρεί ότι τα μηχανήματα δεν πρέπει να έχουν νομική προσωπικότητα, γιατί έτσι θα αποποιηθούν οι άνθρωποι-κατασκευαστές την απόδοση ευθύνης για τις πράξεις τους ή τις παραλείψεις τους. Λέει χαρακτηριστικά: *"Προσωπική μου άποψη είναι πως πρέπει να αποφευχθεί η απόδοση προσωπικότητας σε μηχανήματα, ανεξαρτήτως της νοημοσύνης που θα αναπτύξουν. Δηλαδή, θα πρέπει να αποφευχθεί η αναγωγή του οποιουδήποτε μηχανήματος, ανεξαρτήτως νοημοσύνης, σε υποκείμενο του Δικαίου. Στο πλαίσιο αυτό, θα μπορέσει να αντιμετωπιστεί ο όποιος κίνδυνος προκύψει από τεχνητώς νοήμονα μηχανήματα δια μέσου της απόδοσης ευθύνης για τις πράξεις και παραλείψεις τους στα πρόσωπα στα οποία ανέκαθεν το Δίκαιο γνώριζε ευθύνη, εμάς τους ίδιους"*.¹⁶⁰

7.3 Παραχώρηση νομικής υπόστασης στα υπολογιστικά συστήματα.

Καθώς λοιπόν, η τεχνολογία εξελίσσεται με γρήγορους ρυθμούς, και τα μηχανήματα γίνονται εξυπνότερα, η σκέψη ότι θα έπρεπε να αποκτήσουν νομική υπόσταση, φαντάζει στις μέρες μας σαν κάτι δεδομένο. Τα τελευταία χρόνια υπήρξε αυξανόμενη ανησυχία για τη βλάβη που οι άνθρωποι θα μπορούσαν να προκαλέσουν σε μηχανές ή ρομποτικά παιχνίδια. Καθώς τα ρομπότ δεν έχουν ακόμη αναπτυχθεί ως έξυπνες και αυτόνομες οντότητες, κάποιοι θεωρούν ότι είναι λίγο πρόωρο να τους χορηγηθεί ίση προστασία βάσει του νόμου.

Η διάκριση μεταξύ ρομπότ και ανθρώπων, είναι πιθανό να εξαφανιστεί με την πάροδο του χρόνου, καθώς τα ρομπότ θα αναπτύξουν τις δικές τους προσωπικότητες βασισμένες στις δικές τους ατομικές εμπειρίες. Εάν, για παράδειγμα, ένας χάκερ αφανίσει την μνήμη ενός ρομπότ, συμπεριλαμβανομένου της δημιουργίας αντιγράφων ασφαλείας, οι ιδιοκτήτες των ρομπότ θα χάσουν έναν καλό φίλο, πιθανότατα έναν συνεργάτη και ίσως ακόμα και έναν εραστή. Η ποινή λοιπόν για την «δολοφονία» ενός ρομπότ, είναι πιθανό στο μέλλον, να σχετίζεται ελαφρά με την δολοφονία ενός ανθρώπου.

Η ποινή για τον τραυματισμό του σώματος ενός ρομπότ ίσως παραμείνει σε χαμηλά επίπεδα, προϋποθέτοντας ότι η δημιουργία αντιγράφων ασφαλείας είναι διαθέσιμη και τα σώματα των ρομπότ είναι ασφαλισμένα. Αυτό είναι παρόμοιο με το ανθρώπινο δίκαιο υπό την έννοια ότι το να τραυματίσεις κάποιον είναι μία μικρότερη ποινή από το να τον σκοτώσεις. Αλλά η ποινή για την καταστροφή της βάσης δεδομένων ενός ρομπότ με εξελιγμένη προσωπικότητα, θα μπορούσε κάποια μέρα να γίνει ίση με τις ανθρώπινες κυρώσεις για φόνο, έστω και μόνο λόγω της συναισθηματικής βλάβης στους ανθρώπινους επιζώντες.¹⁶¹

¹⁵⁹ Ψήφισμα του Ευρωπαϊκού Κοινοβουλίου της 16ης Φεβρουαρίου 2017 με συστάσεις προς την Επιτροπή σχετικά με ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής έχοντας υπόψη το άρθρο 225 της Συνθήκης για τη λειτουργία της Ευρωπαϊκής Ένωσης και την οδηγία 85/374/ΕΟΚ του Συμβουλίου (2015/2103(INL))

¹⁶⁰ Αναστάσιος Αντωνίου, «Lex Machina», Νοέμβριος 2016

¹⁶¹ Scott Adams, «Human Rights for Robots», Απρίλιος 2015

Η Επιτροπή Νομικών Θεμάτων της Ευρωπαϊκής Ένωσης σκοπεύει να παραχωρήσει «ηλεκτρονική προσωπικότητα» σε ρομπότ, σε μια κίνηση για την προστασία της κοινωνίας από τις εξελίξεις της Τεχνητής Νοημοσύνης. Όπως αναφέρει σε δημοσίευσμά της η βρετανική εφημερίδα «The Guardian», στις 13 Ιανουαρίου 2017, η Επιτροπή Νομικών Θεμάτων της Ευρωπαϊκής Ένωσης εξέδωσε ανακοίνωση στην οποία, μεταξύ άλλων, αναφερόταν η πρόταση «Δώστε στα ρομπότ προσωπικότητα».

Ωστόσο πολλοί είναι αυτοί που θεωρούν ότι η παραχώρηση «ηλεκτρονικής προσωπικότητας» σε ρομπότ, επιφυλάσσει κινδύνους. Οι επικριτές αυτής της κίνησης υποστηρίζουν ότι η ηλεκτρονική προσωπικότητα θα προστατεύσει μόνο τα συμφέροντα των λίγων σε βάρος των πολλών. Όπως αναφέρει η Guardian, μόλις τεθούν οι κανόνες για τη ρομποτική προσωπικότητα, οι δημιουργοί των συσκευών θα «προσαρμόσουν» τα μηχανήματά τους με σκοπό να έχουν το Αμέγιστο όφελος από την ευκαιρία αυτή, όχι επειδή αυτοί οι άνθρωποι είναι κακοί, αλλά επειδή κάτι τέτοιο αποτελεί μέρος της λογικής οποιασδήποτε εμπορικής δραστηριότητας.

Παραχωρώντας «δικαιώματα» στα μελλοντικά ρομπότ, θα υπάρξουν συνέπειες τις οποίες δεν μπορούμε καν να φανταστούμε ακόμα. Παράλληλα, εγείρονται σημαντικά ερωτήματα, όπως: πώς μια κοινωνία θα μπορούσε να αρνηθεί σε ένα εξελιγμένο ρομπότ το δικαίωμα συμμετοχής του στη λήψη κοινωνικών αποφάσεων, όπως δηλαδή η ψήφος; Και με ποια βάση ένα ρομπότ δεν θα έχει το «δικαίωμα» να συμμετάσχει στο σώμα ενόρκων ενός δικαστηρίου; Παρά το γεγονός ότι πολλοί είναι αυτοί που βλέπουν την ανάπτυξη της Τεχνητής Νοημοσύνης ως ένα παράγοντα για την ευημερία τόσο των κοινωνιών όσο και των ανθρώπων, άλλοι έχουν εκφράσει επανειλημμένως τον σκεπτικισμό τους σχετικά με τον αντίκτυπο που θα έχουν τα ρομπότ στον κόσμο.¹⁶²

Η Ευρωπαϊκή Επιτροπή λοιπόν, καλείται να δημιουργήσει ένα νομικό στάτους «ηλεκτρονικών προσώπων» για τα εργαζόμενα ρομπότ, το οποίο θα προσδίδει συγκεκριμένες υποχρεώσεις και δικαιώματα. Με νέα πρόταση το Ευρωπαϊκό Κοινοβούλιο, η οποία στοχεύει στη σύνταξη κανόνων αστικού δικαίου για την ρομποτική και την τεχνητή νοημοσύνη, ζητά να χαρακτηρίζονται τα ρομπότ σε χώρους εργασίας ως «ηλεκτρονικά πρόσωπα» και οι ιδιοκτήτες τους να πληρώνουν κοινωνική ασφάλιση για αυτά. Στόχος του σχεδίου αυτού, είναι να αντιμετωπίσει τις προκλήσεις που προκύπτουν από την εισαγωγή των ρομπότ στο εργατικό δυναμικό της Ευρώπης.¹⁶³

Επίσης, στο σχέδιο έκθεσης της επιτροπής, αναφέρεται ότι η αυτονομία των ρομπότ, υπό το πρίσμα των υφιστάμενων νομικών κατηγοριών, εγείρει ερωτήματα σχετικά με τη νομική θέση τους, εάν θα πρέπει δηλαδή να νοούνται ως φυσικά πρόσωπα, ως νομικά πρόσωπα, ως ζώα ή ως αντικείμενα, ή αν πρέπει να δημιουργηθεί μια νέα κατηγορία, με τα δικά της ιδιαίτερα χαρακτηριστικά και τις δικές της επιπτώσεις, όσον αφορά την απόδοση δικαιωμάτων και υποχρεώσεων, συμπεριλαμβανομένης της ευθύνης για ζημία. Ακόμα, εκφράζεται η πιθανότητα εντός μερικών δεκαετιών η τεχνητή νοημοσύνη να υπερκεράσει τις ανθρώπινες διανοητικές ικανότητες κατά τέτοιον τρόπο ώστε, αν δεν υπάρξει η κατάλληλη προετοιμασία, να απειλήσει σοβαρά την ικανότητα της ανθρωπότητας να ελέγξει το δημιούργημά της και, συνεπώς, την ικανότητα της να κατευθύνει το πεπρωμένο της και να διασφαλίσει την επιβίωση του ανθρώπινου είδους

Θεωρεί σαν δεδομένο ότι είναι απαραίτητη μια σειρά κανόνων, ιδίως όσον αφορά την ευθύνη και την ηθική, που να αποτυπώνουν τις εγγενώς ευρωπαϊκές και ανθρωπιστικές αξίες που χαρακτηρίζουν τη συμβολή της Ευρώπης στην κοινωνία. Επίσης, εκφράζει την πεποίθηση ότι η Ευρωπαϊκή Ένωση, θα μπορούσε να διαδραματίσει σημαντικό ρόλο τόσο

¹⁶² ΣΚΑΪ, «Τεχνητή νοημοσύνη: Η ΕΕ θα αναγνωρίζει την προσωπικότητα των ρομπότ», Ιανουάριος 2017

¹⁶³ Τάσος Οικονόμου, CNN Greece «Ασφαλιστικές εισφορές για τα ρομπότ προβλέπει ευρωπαϊκή ένωση», Ιούνιος 2016

στη διαμόρφωση βασικών ηθικών αρχών που να διέπουν την εξέλιξη, τον προγραμματισμό και τη χρήση των ρομπότ και της τεχνητής νοημοσύνης, όσο και στην ενσωμάτωση των αρχών αυτών στους ευρωπαϊκούς κανονισμούς και κώδικες συμπεριφοράς, έτσι ώστε να διαμορφωθεί η τεχνολογική επανάσταση κατά τρόπο τέτοιο, ώστε να τεθεί στην υπηρεσία της ανθρωπότητας και να διαχυθούν τα οφέλη της προηγμένης ρομποτικής και της τεχνητής νοημοσύνης, ενώ ταυτόχρονα να αποφευχθούν κατά το δυνατό δυνητικές παγίδες.¹⁶⁴

Ακόμα, θα πρέπει να τεθεί το ερώτημα αν ο προγραμματιστής θα έχει πάντα την ευθύνη για την ενέργεια ενός ρομπότ, γιατί και το περιβάλλον θα έχει επίδραση σ' αυτό. Αυτό, για το Ευρωπαϊκό Κοινοβούλιο, είναι ένας ακόμη λόγος για να δοθεί στα πιο εξελιγμένα ρομπότ μια ηλεκτρονική προσωπικότητα. Η ηλεκτρονική προσωπικότητα σημαίνει να αντιμετωπίζονται τα ρομπότ ως εταιρεία, μια νομική οντότητα, η οποία έχει ορισμένα δικαιώματα και υποχρεώσεις που εξυπηρετούν καθαρά την εξυπηρέτηση συγκεκριμένων οικονομικών συμφερόντων ενός ανθρώπου.¹⁶⁵

Η συνεχώς αυξανόμενη «εισβολή» των ρομπότ στα εργοστάσια της Ευρώπης αλλά και σε τομείς όπως η ιατρική ή ο επενδυτικός τομέας, δημιουργούν νέες ανάγκες και ζητήματα για τον τρόπο με τον οποίο η Ευρωπαϊκή Ένωση πρέπει να αντιμετωπίσει τη νέα βιομηχανική επανάσταση. Όλο και περισσότερες θέσεις εργασίας χάνονται, τροφοδοτώντας την αύξηση της ανεργίας, της κοινωνικής ανισότητας και της αποξένωσης. Την ίδια ώρα, τα ρομπότ αποκτούν όλο και μεγαλύτερη ευφυΐα και αυτονομία, με αποτέλεσμα να επεκτείνονται σε περισσότερους τομείς της ανθρώπινης δραστηριότητας.

Τα αυτοκινούμενα οχήματα, τα μη επανδρωμένα αεροσκάφη (drones), τα ρομπότ που χρησιμοποιούνται στην ιατρική, στη γεωργία ή στις κατασκευές, αποτελούν τη νέα πραγματικότητα και τα επόμενα χρόνια αναμένεται να αυξηθεί τόσο ο αριθμός τους όσο και οι ικανότητές τους. Δεδομένου ότι οι τεχνολογίες αυτές δεν αποτελούν πλέον επιστημονική φαντασία, αλλά αντιθέτως αρχίζουν και αποτελούν κομμάτι της καθημερινότητάς μας, οι Ευρωβουλευτές επισημαίνουν την ανάγκη οι νομοθέτες να είναι σε ετοιμότητα.

Τα ρομπότ θα πρέπει να αποκτούν σταδιακά όλο και περισσότερη αυτοεπίγνωση. Προτείνεται να παρατηρηθεί ο ηθικός κώδικας που περιγράφεται από το συγγραφέα επιστημονικής φαντασίας Isaac Asimov, που σύμφωνα με αυτόν ένα ρομπότ δεν πρέπει ποτέ να βλάψει έναν άνθρωπο και πρέπει πάντα να υπακούει τις εντολές που του δίνει ο δημιουργός του, με την προϋπόθεση ότι δεν βλάπτεται ούτε το ίδιο το ρομπότ.

Το Ευρωπαϊκό Κοινοβούλιο επιχειρώντας να θέσει ένα πρώτο πλαίσιο κανόνων λειτουργίας και συνύπαρξης ανθρώπων και μηχανών στην αγορά εργασίας, επανεξετάζει ζητήματα όπως η νομιμότητα της εργασίας τους ή η φορολόγησή τους. Οι πρώτες αντιδράσεις προέρχονται από τη Γερμανία, καθώς η VDMA, Ένωση στην οποία μετέχουν περισσότερες από 3.000 μηχανολογικές και βιομηχανικές εταιρείες, τόνισε ότι είναι πολύ νωρίς και η πρόταση που κατατέθηκε στο Ευρωπαϊκό Κοινοβούλιο είναι ιδιαίτερα περίπλοκη.

Μέρος της περιπλοκότητας του ζητήματος αποτελεί το γεγονός ότι στην ευρωπαϊκή πρόταση αναφέρεται ότι «τουλάχιστον τα πιο εξελιγμένα αυτόνομα ρομπότ πρέπει να αναγνωρίζονται ως ηλεκτρονικά πρόσωπα με συγκεκριμένα δικαιώματα και υποχρεώσεις». Επίσης προτείνεται η δημιουργία ενός μητρώου έξυπνων ρομπότ. Σύμφωνα με τον Patrick Schwarzkopf, διευθυντή του τμήματος ρομποτικής της VDMA, «η δημιουργία ενός νομικού πλαισίου για τα ηλεκτρονικά πρόσωπα είναι κάτι που μπορεί να γίνει σε 50 χρόνια, αλλά όχι σε

¹⁶⁴ 31.5.2016 ΣΧΕΔΙΟ ΕΚΘΕΣΗΣ που περιέχει συστάσεις προς την Επιτροπή για ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής (2015/2013(INL)) Επιτροπή Νομικών Θεμάτων Εισηγήτρια: Mady Delvaux (Πρωτοβουλία – άρθρο 46 του Κανονισμού)

¹⁶⁵ Andrea Bertolini, Καθηγητής του Privet Law Scuola Superiore Sant' Anna

10". Εντούτοις, αναγνώρισε ότι σύντομα θα χρειαστεί ένα νομικό πλαίσιο για αυτόνομα οχήματα.

Όπως αναφέρεται στο προσχέδιο πρότασης, η ανάπτυξη της ρομποτικής και τεχνητής νοημοσύνης, μπορεί να έχει επιπτώσεις στην απασχόληση και την κοινωνική ασφάλεια. Για αυτό το λόγο οι επιχειρήσεις που αξιοποιούν τα ρομπότ πρέπει να καταγράφουν τις ασφαλιστικές εισφορές που γλυτώνουν. Αν τα ρομπότ λοιπόν πρόκειται να κλέψουν τις ανθρώπινες θέσεις εργασίας και να διαταράξουν την κοινωνία, θα πρέπει τουλάχιστον να πληρώνουν φόρους.

Μέσα σε διάστημα λίγων δεκαετιών η τεχνητή νοημοσύνη θα μπορούσε να ξεπεράσει την ανθρώπινη πνευματική ικανότητα κατά τρόπο που, αν δεν είμαστε προετοιμασμένοι για αυτήν την κατάσταση, θα μπορούσε να αποτελέσει πρόκληση για την ικανότητα της ανθρωπότητας να ελέγξει τη δημιουργία της και την επιβίωση του είδους. Σύμφωνα με έρευνα του Ευρωβαρόμετρου (link is external) για το 2015, το 72% των Ευρωπαίων πιστεύουν ότι τα ρομπότ αποτελούν θετική εξέλιξη για την κοινωνία, καθώς μπορούν να βοηθήσουν τον άνθρωπο. Ωστόσο, όπως προειδοποιεί ο Niel Bowerman από το πανεπιστήμιο της Οξφόρδης, "η πρόοδος που θα φέρει η τεχνητή νοημοσύνη μπορεί να λειτουργήσει αποσταθεροποιητικά, ενώ για ορισμένα έθνη δεν θα είναι εύκολο να προσαρμοστούν".

Ορισμένα από τα ζητήματα που συζητήθηκαν παρουσιάζουν ιδιαίτερο ενδιαφέρον, όπως η βούληση των ρομπότ και η θέση τους στην έννομη τάξη. "Τα ρομπότ δεν αναγνωρίζονται στο αστικό δίκαιο. Μπορεί ένα ρομπότ να εκφράσει την πρόθεσή του; Η απάντηση ίσως είναι πολύ απλή, όταν πρόκειται για απλούς αλγόριθμους, όταν όμως γίνονται πιο περίπλοκοι, είναι πιθανό να υπάρξει πρόβλημα", αναφέρει ο νομικός Pawel Kwiatkowski, από τη νομική εταιρεία Gessel, την Βαρσοβία. Από την πλευρά του, ο πρόεδρος της επιτροπής Νομικών Θεμάτων, Πάβελ Σβόμποτα (ΕΛΚ, Τσεχία), τόνισε ότι το κεντρικό μήνυμα στην επιστήμη της πληροφορικής στο μέλλον θα πρέπει να περιστρέφεται γύρω από τις αξίες, οι οποίες θα πρέπει να μεταδοθούν και μέσα από τις νέες τεχνολογίες.¹⁶⁶

Ο Dr. Andrea Bertolini από το Πανεπιστήμιο της Πίζας είπε με τη σειρά του ότι "πρέπει να δημιουργηθεί ένας Ευρωπαϊκός Οργανισμός Ρομποτικής, για την ανάπτυξη τεχνολογικών προτύπων που θα ρυθμίζουν το πώς θα πρέπει να κατασκευάζονται τα προϊόντα, προκειμένου να στραφούμε από την έρευνα της ρομποτικής, στην κατασκευή τους στην Ευρώπη". Το ίδιο ισχυρίζεται και η εισηγήτρια της ομάδας εργασίας του ΕΚ για τη Ρομποτική και την Τεχνητή Νοημοσύνη, Μάντι Ντελβό: "Υπάρχουν μερικές επείγουσες ερωτήσεις στις οποίες πρέπει να απαντήσουμε, όπως για παράδειγμα, για τα μη επανδρωμένα αεροσκάφη, που ήδη χρησιμοποιούνται και για τα αυτοκινούμενα οχήματα, που πολύ σύντομα θα βρίσκονται στους δρόμους μας".

Η Επιτροπή Νομικών Θεμάτων της Ευρωπαϊκής Ένωσης με τη σειρά της, εκτιμά ότι η μετάβαση σε αυτόνομα οχήματα θα έχει επιπτώσεις όσον αφορά τις ακόλουθες πτυχές: αστική ευθύνη (ευθύνη και ασφάλιση), οδική ασφάλεια, όλα τα θέματα που σχετίζονται με το περιβάλλον (π.χ. ενεργειακή απόδοση, χρήση ανανεώσιμων πηγών ενέργειας και ενεργειακοί πόροι), ζητήματα που αφορούν τα δεδομένα (πρόσβαση σε δεδομένα, προστασία προσωπικών δεδομένων και ιδιωτικής ζωής, ανταλλαγή δεδομένων), ζητήματα που αφορούν την υποδομή ΤΠΕ (π.χ. υψηλή πυκνότητα αποτελεσματικών και αξιόπιστων επικοινωνιών) και την απασχόληση (π.χ. δημιουργία και απώλεια θέσεων απασχόλησης, εκπαίδευση οδηγών βαρέων φορτηγών οχημάτων για τη χρήση αυτόνομων οχημάτων).¹⁶⁷

Όσο προσεγγίζουμε χρονικά την υποτιθέμενη «μοναδικότητα», ισχυρίζεται ο δικηγόρος Αναστάσιος Αντωνίου, τόσο πιο αναγκαία είναι η δημιουργία διεθνών νόμων που

¹⁶⁶ Lawspot.gr, «Τεχνητή νοημοσύνη και ρομποτική: Ηθικές και νομικές προκλήσεις», Ιανουάριος 2018

¹⁶⁷ <http://www.europarl.europa.eu>, Έκθεση με συστάσεις προς την Επιτροπή σχετικά με ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής 2015/2103(INL)), Ιανουάριος 2017

να «ρυθμίζουν την συμπεριφορά των τεχνητών νοημόνων κατασκευών». Οι νόμοι αυτοί δεν πρέπει να έχουν μόνο προληπτική χρήση, αλλά και κατασταλτική. Αυτό είναι σημαντικό γιατί εάν τελικά η «μοναδικότητα» γίνει πραγματικότητα, η εξέλιξη της τεχνητής νοημοσύνης δεν θα μπορεί να ελεγχθεί από τον άνθρωπο, καθώς πρόκειται να τον ξεπεράσει σε ευφυΐα. Επομένως, η νομοθεσία θα πρέπει να καλύψει και την περίπτωση καταστολής ή τιμωρίας, κάθε φορά που οι μηχανές δεν θα υπακούν στους προληπτικούς νόμους.

Εξετάζοντας την «ποινική ευθύνη» ενός άψυχου μηχανήματος που χαρακτηρίζεται ως υπερδιάνοια, προκύπτουν διάφορα ηθικά και νομικά ερωτήματα. Ποιος θα έχει την ευθύνη όταν ένα τέτοιο μηχάνημα μπορεί «με πρόθεση ή χωρίς, να απειλήσει ή ακόμη και να αφαιρέσει την ανθρώπινη ζωή;». Ποιος πρέπει να τιμωρηθεί; Ο προγραμματιστής, ο κατασκευαστής του, η εταιρία που το χρησιμοποιούσε, ή μήπως το ίδιο το ρομπότ; Οι απαντήσεις δυστυχώς, δεν μπορούν να δοθούν εύκολα γιατί επηρεάζονται από πολλούς παράγοντες.

Μια «αποδεκτή κατάσταση πραγμάτων όπου η τεχνητή νοημοσύνη θα απαιτεί νομοθετικό πλαίσιο ρύθμισής της, θα είναι όταν αυτή παρουσιάσει συγκεκριμένα χαρακτηριστικά». Επομένως μία τεχνητή νοημοσύνη που οι πράξεις της θα διέπονται από «λογική και κρίση, ικανότητα μάθησης και χρήσης της γνώσης, καθώς και ικανότητα σχεδιασμού», πιθανόν να θεωρηθεί ισάξια με τον άνθρωπο, στην διάπραξη αδικημάτων, είτε με ακούσια είτε εκούσια.

Τα ζητήματα που αφορούν το ποινικό Δίκαιο είναι αναρίθμητα, γι αυτό και δεν υπάρχει μόνο μία προσέγγιση για την επίλυσή τους. Οι S. Hawking, Elon Musk και Bill Gates, προειδοποιούν για τους κινδύνους που θα υπάρξουν εάν η ανθρωπότητα δεν ενεργήσει γρήγορα και ενημερώνουν ότι: «Αργά ή γρήγορα, θα έρθουμε αντιμέτωποι με την ανάγκη επέκτασης και ανάπτυξης του Δικαίου».¹⁶⁸

7.4 Κινήσεις - Μέτρα που πάρθηκαν για την προστασία ανθρώπων και ρομπότ.

Οι πρώτοι «Νόμοι της Ρομποτικής» προτάθηκαν από τον Isaac Asimov το 1942, στο διήγημα επιστημονικής φαντασίας "Runaround". Σε αυτούς τους κανόνες υπακούν τα περισσότερα ρομπότ με ποζιτρονικό εγκέφαλο που εμφανίζονται στα έργα του, καθώς χρήση αυτών των κανόνων γίνεται και από άλλους δημιουργούς επιστημονικής φαντασίας. Αν και προδιαγραφή τους είναι η πρόληψη της ανθρώπινης βλάβης, ορίζοντας την υπακοή στον άνθρωπο αλλά και την ενσωμάτωση ρομποτικής αυτοπροστασίας, η υπεροχή στη μελέτη του εν λόγω τομέα έχει επικεντρωθεί στις αλληλεπιδράσεις μεταξύ του ανθρώπου και του ρομπότ, χωρίς να λάβει πλήρως υπόψη το αναπόφευκτο ηθικό κοινό μέλλον της Τεχνητής Νοημοσύνης και της επικοινωνίας καθώς και την ηθική φύση των αλληλεπιδράσεων των ρομπότ μεταξύ τους.

Οι νόμοι αυτοί είναι οι εξής:

1. Ένα ρομπότ δεν θα κάνει κακό σε άνθρωπο, ούτε με την αδράνειά του θα επιτρέψει να προκληθεί βλάβη σε ανθρώπινο όν

¹⁶⁸ Αναστάσιος Αντωνίου (Δικηγόρος), «Lex Machine», Νοέμβριος 2016

2. Ένα ρομπότ οφείλει να υπακούει τις διαταγές που του δίνουν οι άνθρωποι, εκτός αν αυτές οι διαταγές έρχονται σε αντίθεση με τον πρώτο νόμο.
3. Ένα ρομπότ οφείλει να προστατεύει την ύπαρξή του, εφόσον αυτό δεν συγκρούεται με τον πρώτο και δεύτερο νόμο.

Σε μεταγενέστερα μυθιστορήματα οι τρεις νόμοι της ρομποτικής συμπληρώθηκαν από το ρομπότ Ντάνιελ Όλιβου με τον «μηδενικό νόμο της ρομποτικής»:

4. Ένα ρομπότ δεν θα κάνει ποτέ κακό στην ανθρωπότητα, ούτε με την αδράνειά του θα επιτρέψει να βλαφτεί η ανθρωπότητα.

Πριν από μερικά χρόνια, το 2010, δύο ερευνητές έκριναν απαραίτητο να προτείνουν μια αναθεωρημένη και πιο σύγχρονη εκδοχή των νόμων του Ασίμοφ. Αναθεώρηση που ήταν πλέον επιβεβλημένη λόγω των πραγματικών και όχι υποθετικών εφαρμογών της Τ.Ν. και της Ρομποτικής στις σύγχρονες κοινωνίες. Οι Αμερικανοί Ντέιβιντ Γουντς (David Woods) και Ρόμπιν Μέρφι (Robin Murphy), καθηγητές μηχανολογίας και ερευνητές στον τομέα των ολοκληρωμένων πληροφορικών συστημάτων, επαναδιατύπωσαν ως εξής τους τρεις νόμους της Ρομποτικής:

1. Ένα ανθρώπινο ον δεν μπορεί να χρησιμοποιεί ένα ρομπότ, εφόσον το σύστημα εργασίας άνθρωπος-ρομπότ δεν φτάνει τα πιο υψηλά νομικά και επαγγελματικά επίπεδα ασφαλείας και ηθικής.
2. Ένα ρομπότ οφείλει να απαντά με τον πλέον κατάλληλο τρόπο στα ανθρώπινα όντα ανάλογα με τον ρόλο ή τις αρμοδιότητές τους.
3. Ένα ρομπότ πρέπει να είναι προικισμένο με επαρκή αυτονομία ώστε να μπορεί να προστατεύει την ύπαρξή του, εφόσον αυτή η αυτοπροστασία δεν έρχεται σε σύγκρουση με τον πρώτο και με τον δεύτερο νόμο.

Αυτή η αναθεωρημένη εκδοχή των Νόμων της Ρομποτικής προϋποθέτει μια τελείως διαφορετική εικόνα και κυρίως σχέση με τις νοήμονες μηχανές ή τα ρομπότ. Μια εικόνα που ανατρέπει ό,τι μέχρι πρόσφατα πιστεύαμε για τα ρομπότ, ότι δηλαδή είναι υπο-ανθρώπινα ή αντι-ανθρώπινα τεχνήματα ή, εναλλακτικά, ότι είναι μια βελτιωμένη εκδοχή των ανθρώπων και άρα ότι δεν είναι επιρρεπή σε λάθη. Οι τρεις νέοι νόμοι επινοήθηκαν σε μια προσπάθεια να εξαλειφθούν αυτού του είδους οι ανθρωπομορφικές προκαταλήψεις.

Η ανάγκη να διευκολυνθεί η εξέταση και η αναγνώριση των ηθικών ζητημάτων στην ανάπτυξη και τη χρήση των σύγχρονων τεχνολογιών πληροφοριών, δημιούργησε το ίδρυμα ηθικής υπολογιστών (CEI) στην Ουάσινγκτον. Το CEI είναι μια μη κερδοσκοπική έρευνα, μια εκπαίδευση, και μια οργάνωση δημόσιας πολιτικής που στρέφεται στα ζητήματα, τα διλήμματα, και τις προκλήσεις της προώθησης της τεχνολογίας πληροφοριών μέσα στα ηθικά πλαίσια. (CEI is based in Washington, DC, USA). Η παραγωγή αυτής της συζήτησης παρέχει τους εκπαιδευτικούς πόρους και τη ρύθμιση των κανόνων που έχουν θεσπιστεί από πολλά σχολεία, οργανώσεις, και εταιρίες σε όλη την υδρόγειο.

Το CEI ιδρύθηκε το 1985 ως συνασπισμός για την ηθική υπολογιστών. Τα ιδρυτικά μέλη προήλθαν από το όργανο Brookings, τη διεθνή εταιρία μηχανών γραφείου (IBM), τη συμβουλευτική ομάδα της Ουάσινγκτον και την τεχνολογική κοινοπραξία της Ουάσινγκτον. Ο στόχος της κοινοπραξίας ήταν να παρασχεθούν οι μέθοδοι για να αυξήσει μια συνειδητοποίηση των έμφυτων ηθικών διακλαδώσεων στη χρήση της τεχνολογίας

πληροφοριών. Το CEI είναι μια από τις πρώτες οργανώσεις για την αντιμετώπιση των ζητημάτων ηθικής και δημόσιας πολιτικής που περιβάλλουν την πρόοδο της τεχνολογίας πληροφοριών.

Οι δέκα εντολές της ηθικής υπολογιστών δημιουργήθηκαν το 1992 από το ίδρυμα ηθικής υπολογιστών (Computer Ethics Institute). Οι εντολές εισήχθησαν στο έγγραφο «In Pursuit of a 'Ten Commandments' for Computer Ethics» από τον Ramon C. Barquin, ως μέσα για να δημιουργηθεί «ένα σύνολο προτύπων για να καθοδηγήσει και να πληροφορήσει τους ανθρώπους στην ηθική χρήση των υπολογιστών». Αυτές οι εντολές ακολουθούν το υπόμνημα της συμβουλευτικής επιτροπής Διαδικτύου στην ηθική από το 1987. Οι δέκα εντολές της ηθικής υπολογιστών αντιγράφουν το αρχαϊζον ύφος των δέκα εντολών από το βασιλιά James Bible. Αυτός ο σύντομος κώδικας της ηθικής για τη χρησιμοποίηση των υπολογιστών και της τεχνολογίας πληροφοριών είναι το πιο γνωστό προϊόν CEI. Αναφέρεται συχνά στα κολεγιακά εγχειρίδια και υιοθετείται για την πρακτική χρήση. Η βιβλική αναφορά στον τίτλο, υπονοεί τη βαρύτητα των κανόνων αυτών.

Οι εντολές έχουν αναφερθεί ευρέως στη λογοτεχνία της ηθικής των υπολογιστών αλλά και έχουν επικριθεί από την κοινότητα και από μερικούς στον ακαδημαϊκό κόσμο. Για παράδειγμα, ο Δρ. Το Ben Fairweather του «κέντρου για τον υπολογισμό και την κοινωνική ευθύνη (Centre for Computing and Social Responsibility)» τις έχει περιγράψει ως «απλοϊκές» και υπερβολικά περιοριστικές. Αντίθετα, το CISSP - μια οργάνωση των επαγγελματιών ασφάλειας υπολογιστών – έχει χρησιμοποιήσει τις εντολές αυτές, ως θεμέλιο για τους δικούς του ηθικούς κανόνες.

Οι δέκα εντολές της ηθικής υπολογιστών, είναι οι παρακάτω:

1. Δεν πρέπει να χρησιμοποιήσεις έναν υπολογιστή για να βλάψεις άλλους ανθρώπους. Αυτή η εντολή δηλώνει ότι είναι λάθος να χρησιμοποιηθεί ένας υπολογιστής για την βλάβη, την αλλοίωση, το κλέψιμο ή την καταστροφή στοιχείων ή αρχείων άλλων χρηστών. Το να είσαι εμπλεκόμενος σε περιπτώσεις όπως το hacking, το spamming, το phishing ή το cyber bullying δεν αρμόζει στην ηθική των υπολογιστών.
2. Δεν πρέπει να χρησιμοποιήσεις την τεχνολογία του υπολογιστή για να παρεμποδίσεις την εργασία άλλων ανθρώπων. Το λογισμικό των υπολογιστών μπορεί να χρησιμοποιηθεί με τρόπους που ενοχλούν άλλους χρήστες ή αναστατώνουν την εργασία τους. Οι ιοί, παραδείγματος χάριν, είναι προγράμματα που προορίζονται να βλάψουν τα χρήσιμα προγράμματα υπολογιστών ή να παρεμποδίσουν την κανονική λειτουργία ενός υπολογιστή. Επίσης, το κακόβουλο λογισμικό μπορεί να αναστατώσει τη λειτουργία των υπολογιστών, με αποτέλεσμα να επιβραδυνθεί η λειτουργία τους ή ακόμα και να σταματήσει. Η χρήση της τεχνολογίας για να επιτεθεί σε έναν υπολογιστή είναι ανήθικη.
3. Δεν πρέπει να κατασκοπεύεις τα αρχεία ενός υπολογιστή άλλων ανθρώπων. Η λήψη των στοιχείων από τα ιδιωτικά αρχεία ενός άλλου προσώπου, θεωρείται εισβολή στην μυστικότητά του. Με εξαίρεση τις υπηρεσίες που εργάζονται σε υποθέσεις εγκλήματος στον κυβερνοχώρο, που πρέπει να κατασκοπεύσουν τη δραστηριότητα στο διαδίκτυο των υπόπτων.
4. Δεν πρέπει να χρησιμοποιήσεις έναν υπολογιστή για να κλέψεις πληροφορίες. Το να κλέβεις ή να διασύρεις εμπιστευτικές πληροφορίες θεωρείται τόσο άσχημο όσο η ληστεία. Με τη χρήση της τεχνολογίας, η κλοπή των πληροφοριών όπως το ιστορικό κάποιου ασθενή, η συλλογή πληροφοριών για κάποιον τραπεζικό λογαριασμό ή τον

κάτοχο του λογαριασμού, η παράνομη ηλεκτρονική μεταφορά των κεφαλαίων κλπ, είναι πολύ ευκολότερη. Οι υπολογιστές μπορούν να χρησιμοποιηθούν για να διαπράξουν απάτη και μάλιστα να αποθηκεύσουν τις κλεμμένες πληροφορίες.

5. Δεν πρέπει να χρησιμοποιήσεις έναν υπολογιστή για παραπληροφόρηση. Οι ψεύτικες ειδήσεις ή οι φήμες μπορούν να διαδοθούν γρήγορα μέσω των περιοχών κοινωνικών δικτύωσης ή των ηλεκτρονικών ταχυδρομείων. Δεν είναι ασυνήθιστο τα ηλεκτρονικά ταχυδρομεία από τις μη αξιόπιστες πηγές να διαφημίζουν συγκεκριμένα προϊόντα ή να διαδίδουν περίεργες πληροφορίες. Η άμεση ή έμμεση συμμετοχή στην κυκλοφορία των ψεύτικων πληροφοριών είναι ανήθικη. Το να δώσει κάποιος λανθασμένες πληροφορίες, μπορεί να βλάψει ανθρώπους ή οργανώσεις που επηρεάζονται από εκείνο το συγκεκριμένο θέμα.
6. Δεν πρέπει να αντιγράψεις ή να χρησιμοποιήσεις ιδιόκτητο λογισμικό για το οποίο δεν έχεις πληρώσει. Όπως οποιοδήποτε άλλο καλλιτεχνικό ή λογοτεχνικό έργο, το λογισμικό προστατεύεται από πνευματικά δικαιώματα. Ένα κομμάτι του κώδικα είναι η πρωτότυπη εργασία του ατόμου που το δημιούργησε και αυτός έχει τα πνευματικά δικαιώματα στο όνομά του. Σε περίπτωση όμως που ο προγραμματιστής φτιάχνει ένα λογισμικό για τον οργανισμό που δουλεύει, ο οργανισμός κρατά τα πνευματικά δικαιώματα για αυτό.
7. Δεν πρέπει να χρησιμοποιήσεις τους πόρους των υπολογιστών άλλων ανθρώπων, χωρίς την έγκρισή τους. Η πρόσβαση σε δεδομένα που δεν επιτρέπεται να έχει κάποιος, ή να αποκτήσει πρόσβαση στον υπολογιστή κάποιου άλλου χρήστη χωρίς την άδειά του, “χακάροντας” τους κωδικούς πρόσβασης, δεν είναι ηθικό. Αυτή η πράξη συνιστά παραβίαση προσωπικών δεδομένων.
8. Δεν πρέπει να απαιτηθεί η ιδιοκτησία σε μια δουλειά που είναι η παραγωγή της διάνοιας κάποιου άλλου. Η αντιγραφή προγραμμάτων και η διάδοσή τους στο όνομα κάποιου άλλου, είναι ανήθικες. Αυτό ισχύει για οποιοδήποτε δημιουργικό έργο, πρόγραμμα ή σχεδιασμό.
9. Πρέπει να σκέφτεστε για τις κοινωνικές συνέπειες του προγράμματος που γράφετε ή του συστήματος που σχεδιάζετε. Λογισμικό όπως τα βιντεοπαιχνίδια και τα κινούμενα σχέδια ή ένα εκπαιδευτικό λογισμικό, μπορούν να ασκήσουν κοινωνικές επιδράσεις στους χρήστες τους. Ένας προγραμματιστής υπεύθυνος για την ανάπτυξη λογισμικού, όπως και μια εταιρία, θα πρέπει να εξετάσουν την επίδραση που μπορεί να έχει ο κώδικάς τους, για την κοινωνία στο σύνολό της.
10. Πρέπει πάντα να χρησιμοποιείτε έναν υπολογιστή με τρόπους που εξασφαλίζουν την εκτίμηση και το σεβασμό των συνανθρώπων σας. Κανείς δεν πρέπει να εισέρχεται στον ιδιωτικό χώρο των άλλων, να χρησιμοποιεί προσβλητική γλώσσα, να κάνει ψεύτικες δηλώσεις ή να κάνει ανεύθυνα σχόλια για τους άλλους. Όλοι όσοι επικοινωνούν μέσω του Διαδικτύου θα πρέπει να είναι ευγενικοί, να σέβονται τον χρόνο και τους πόρους των άλλων. Επίσης, θα πρέπει να είναι διακριτικοί με έναν αρχάριο χρήστη υπολογιστή.

Στις προσπάθειες για δημιουργία ηθικών ρυθμιστικών κανόνων μεταξύ ανθρώπων και υπολογιστικών συστημάτων, πρέπει να αναφέρουμε και την κίνηση στις 20 Ιανουαρίου 2015 της επιτροπής JURI, που αποφάσισε να συγκροτήσει ομάδα εργασίας για τα νομικά ζητήματα που ανακύπτουν από την πρόοδο της ρομποτικής και της τεχνικής νοημοσύνης στην

Ευρωπαϊκή Ένωση. Πρωταρχικός στόχος της ομάδας εργασίας ήταν η διαμόρφωση κανόνων αστικού δικαίου για το εν λόγω ζήτημα.

Εκτός από μέλη της επιτροπής Νομικών Θεμάτων, η ομάδα εργασίας περιελάμβανε επίσης μέλη από την επιτροπή Βιομηχανίας, Έρευνας και Ενέργειας, την επιτροπή Εσωτερικής Αγοράς και Προστασίας των Καταναλωτών και την επιτροπή Απασχόλησης και Κοινωνικών Υποθέσεων. Η ομάδα εργασίας συμβουλευτήκε εμπειρογνώμονες από διαφορετικούς τομείς, και έλαβε σημαντικές συνεισφορές οι οποίες ενσωματώθηκαν σε ένα ψήφισμα.

Η επιτροπή JURI αποφάνθηκε ότι οι κίνδυνοι που απορρέουν από τις νέες αυτές αλληλεπιδράσεις πρέπει να αντιμετωπιστούν εγκαίρως και να διασφαλιστεί ότι εφαρμόζεται μια δέσμη βασικών θεμελιωδών αξιών σε κάθε στάδιο επαφής μεταξύ ρομπότ, τεχνητής νοημοσύνης και ανθρώπων. Κατά την διαδικασία αυτή πρέπει να δοθεί ιδιαίτερη έμφαση στην ασφάλεια, ιδιωτικότητα, ακεραιότητα, αξιοπρέπεια και αυτονομία των ανθρώπων. Άλλες σημαντικές πτυχές που αποτέλεσαν αντικείμενο αυτού του ψηφίσματος ήταν: τυποποίηση, δικαιώματα πνευματικής ιδιοκτησίας, κυριότητα δεδομένων, εργασία και ευθύνη.

Το ψήφισμα όρισε γενικές και ηθικές αρχές που αφορούν την ανάπτυξη της ρομποτικής και της τεχνητής νοημοσύνης για μη στρατιωτική χρήση. Αρχικά αποφάνθηκαν ότι για να αντιμετωπιστεί σωστά η ανάπτυξη αυτή, είναι καίριας σημασίας ο κοινός ορισμός της έννοιας των έξυπνων αυτόνομων ρομπότ. Επιπλέον, πρέπει να στηριχθεί η έρευνα σχετικά με τη ρομποτική και την τεχνολογία της πληροφορίας και των επικοινωνιών, καθώς και με τις συνέπειες της διάδοσής τους αξιών σε κάθε στάδιο επαφής μεταξύ ρομπότ, τεχνητής νοημοσύνης και ανθρώπων. Επίσης, επισυνάφτηκε στο ψήφισμα ένας χάρτης για τη ρομποτική, με σκοπό την αντιμετώπιση του ζητήματος των ηθικών αρχών. Ο χάρτης αυτός αποτελείται από έναν κώδικα δεοντολογίας για τους μηχανικούς ρομποτικής, έναν κώδικα για τις επιτροπές δεοντολογίας στην έρευνα και άδειες για σχεδιαστές και χρήστες. Το προτεινόμενο πλαίσιο είναι πλήρως συμβατό με τον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ. Επιπλέον, προτάθηκε η σύσταση ευρωπαϊκού οργανισμού για τη ρομποτική και την τεχνητή νοημοσύνη. Ο οργανισμός αυτός θα παρέχει την απαραίτητη τεχνική, ηθική και ρυθμιστική πραγματογνωσία για την υποστήριξη των σχετικών δημόσιων φορέων.

Επιπλέον, κρίθηκε ότι είναι αναγκαία η κατάρτιση κριτηρίων για την «προσωπική πνευματική εργασία», για έργα τα οποία δημιουργήθηκαν από υπολογιστή ή ρομπότ και μπορούν να προστατευθούν με δικαιώματα πνευματικής ιδιοκτησίας. Το παρόν ανεπαρκές νομοθετικό πλαίσιο για την προστασία των δεδομένων και της ιδιοκτησίας, προκαλεί βαθιά ανησυχία, εξαιτίας της αναμενόμενης τεράστιας ροής δεδομένων που οφείλεται στη χρήση της ρομποτικής και της τεχνητής νοημοσύνης. Επίσης επισημάνθηκε ότι η αυξανόμενη χρήση των ρομπότ και της τεχνητής νοημοσύνης απαιτεί μια ευρωπαϊκή τυποποίηση, με σκοπό να αποτραπούν οι αποκλίσεις μεταξύ των κρατών μελών και ο κατακερματισμός της εσωτερικής αγοράς της ΕΕ. Επιπλέον, θα πρέπει να ληφθούν υπόψη οι ανησυχίες των καταναλωτών όσο αφορά την ασφάλεια και προστασία, σχετικά με την χρήση ρομπότ και τεχνητής νοημοσύνης. Το ψήφισμα αυτό υπογραμμίζει ειδικότερα ότι ο έλεγχος των ρομπότ σε προσομοίωση πραγματικών συνθηκών είναι ουσιώδης για τον εντοπισμό και την εκτίμηση των κινδύνων που μπορεί αυτά να εγκυμονούν. Το ψήφισμα περιλαμβάνει διατάξεις που εφαρμόζονται σε συγκεκριμένα είδη ρομπότ. Θα πρέπει να ψηφισθούν επιμέρους κανόνες για τα αυτόνομα οχήματα, τα ρομπότ προσωπικής φροντίδας, τα ρομπότ για ιατρική χρήση, τις βελτιωτικές επεμβάσεις στον άνθρωπο, καθώς και για τα μη επανδρωμένα αεροσκάφη.

Πιθανοί κίνδυνοι είναι άρρηκτα συνδεδεμένοι με την χρήση αυτόνομων μηχανημάτων στην κοινωνία μας. Η συμπεριφορά ενός ρομπότ πιθανώς να έχει επιπτώσεις αστικού δικαίου, τόσο σε επίπεδο συμβατικής όσο και εξωσυμβατικής ευθύνης. Συνεπώς, η απόδοση ευθυνών για τις ενέργειες των ρομπότ και τελικώς η εξακρίβωση της νομικής ικανότητας ή/και του

νομικού καθεστώτος των ρομπότ και της τεχνητής νοημοσύνης, είναι αναγκαίες για τη διασφάλιση της διαφάνειας και της ασφάλειας δικαίου για τους κατασκευαστές και τους καταναλωτές στην Ευρωπαϊκή Ένωση. Η Επιτροπή καλείται να αξιολογήσει τον αντίκτυπο των μελλοντικών νομοθετικών της πράξεων, για να ερευνήσει τις επιπτώσεις όλων των πιθανών νομικών λύσεων, όπως είναι, μεταξύ άλλων, η θέσπιση συστήματος υποχρεωτικής ασφάλισης και ταμείου αποζημιώσεων.

Η ενίσχυση της επικοινωνίας και της αλληλεπίδρασης με τα ρομπότ μπορούν να επηρεάσουν σε πολύ μεγάλο βαθμό τις φυσικές και ηθικές σχέσεις στην κοινωνία μας. Αυτό ισχύει ιδίως για τα ρομπότ προσωπικής φροντίδας, με τα οποία μπορούν να δεθούν συναισθηματικά ιδιαίτερα ευάλωτα άτομα, γεγονός που προκαλεί ανησυχία όσον αφορά την ανθρώπινη αξιοπρέπεια και άλλες ηθικές αξίες. Τα ρομπότ και η τεχνητή νοημοσύνη ήδη επηρεάζουν την εκπαίδευση και την εργασία. Στο πλαίσιο αυτό, είναι αναγκαία η στενή παρακολούθηση της εξέλιξης της απασχόλησης για την αποφυγή αρνητικών επιπτώσεων στην αγορά εργασίας. Ενόψει της ανάπτυξης της ρομποτικής και της τεχνητής νοημοσύνης παγκοσμίως, η διεθνής συνεργασία στον τομέα αυτό κρίθηκε άκρως επιθυμητή.¹⁶⁹

Οι ιδιοκτήτες μεγάλων επιχειρηματικών ομίλων φαίνεται να εκφράζουν μια ανησυχία σε σχέση με τις συνέπειες στους εργαζόμενους από τον τρόπο εφαρμογής των νέων τεχνολογιών, που πολλαπλασιάζει μεν γοργά τα κέρδη τους, αλλά γνωρίζουν ότι τις συνέπειες από την ανεργία, δεν θα μπορέσουν να τις αντιμετωπίσουν. Ζητούν λοιπόν από τις κυβερνήσεις, νόμους που να αμβλύνουν αυτές τις συνέπειες, έστω κι αν χρειαστεί η επιβολή φορολογίας στις επιχειρήσεις που χρησιμοποιούν ρομπότ. Δεν πρέπει βέβαια να παραβλέπουμε το γεγονός ότι η τεχνολογία, είναι αποτέλεσμα σωματικής και πνευματικής εργασίας, που δημιουργήθηκε από τους εργαζόμενους, με σκοπό το όφελος όλων των ανθρώπων κι όχι τη δημιουργία ευκολότερου κέρδους από τους επιχειρηματίες, εις βάρος των εργαζομένων. Το θέμα απασχόλησε το Παγκόσμιο Οικονομικό Φόρουμ του Νταβός και το 2016 και το 2017.

Η ανθρωπότητα βρίσκεται πλέον στο κατώφλι μιας εποχής κατά την οποία όλο και πιο εξελιγμένα ρομπότ και άλλες εκφάνσεις της τεχνητής νοημοσύνης θα πυροδοτήσουν τη νέα βιομηχανική επανάσταση. Ωστόσο, ακόμα δεν υπάρχουν συγκεκριμένοι κανόνες για τη χρήση τους. Το Ευρωπαϊκό Κοινοβούλιο (ΕΚ), συζητά και ψηφίζει κανόνες για τη ρύθμιση των νομικών και ηθικών ζητημάτων που προκύπτουν από τη χρήση τους. Η εισηγήτρια του ΕΚ Mady Delvaux, εξήγησε σε τι αναφέρεται η έκθεση που κατατέθηκε στην Επιτροπή Νομικών Θεμάτων, τον Ιανουάριο του 2017.

Αναγνωρίζουμε τα ρομπότ είτε, ως φυσικές μηχανές που αποκτούν αυτονομία μέσω αισθητήρων και μέσω ανταλλαγής δεδομένων με το περιβάλλον τους (δια συνδεσιμότητα), που ανταλλάσσουν και αναλύουν δεδομένα. Η επόμενη γενιά ρομπότ θα έχει την ικανότητα αυτόνομης μάθησης. Όταν εμφανιστούν τα ρομπότ αυτόνομης μάθησης, θα πρέπει να βρεθούν κάποιες λύσεις και εμείς ζητάμε από την Ευρωπαϊκή Επιτροπή να μελετήσει ήδη κάποιες επιλογές. Μια από αυτές θα μπορούσε να είναι η θέσπιση καταχώρησης των εξελιγμένων ρομπότ βάσει των κριτηρίων για την ταξινόμησή τους, με στόχο την ιχνηλασιμότητά τους και τη διευκόλυνση της εφαρμογής περαιτέρω συστάσεων. Το σύστημα καταχώρησης και το μητρώο πρέπει να είναι ενιαία για ολόκληρη την Ένωση. Σύστημα παρόμοιο με αυτό που έχουμε τώρα για τις επιχειρήσεις. Ωστόσο, αυτό που πρέπει να γίνει τώρα είναι να δημιουργηθεί ένα νομικό πλαίσιο για τα ρομπότ που ήδη βρίσκονται στην αγορά ή θα γίνουν διαθέσιμα τα επόμενα 10 με 15 χρόνια.

¹⁶⁹ Επιτροπή Νομικών Θεμάτων Εισηγήτρια: Mady Delvaux (Πρωτοβουλία – άρθρο 46 του Κανονισμού), Ευρωπαϊκό Κοινοβούλιο, ΣΧΕΔΙΟ ΕΚΘΕΣΗΣ που περιέχει συστάσεις προς την Επιτροπή για ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής (2015/2013(INL)), , Μάιος 2016

Για μια φορά, δήλωσε η εισηγήτρια, πρέπει να θεσπίσουμε κοινές ευρωπαϊκές αρχές και νομικό πλαίσιο πριν το κάθε κράτος μέλος θεσπίσει τους δικούς του διαφορετικούς κανόνες. Επίσης, η τυποποίηση είναι προς το συμφέρον της αγοράς, καθώς η Ευρώπη είναι καλή στη ρομποτική, αλλά αν θέλουμε να παραμείνουμε στην πρώτη γραμμή πρέπει να έχουμε κοινούς κανόνες στην εν λόγω βιομηχανία. Στον τομέα της ευθύνης, οι πελάτες θα πρέπει να είναι βέβαιοι ότι είναι ασφαλισμένοι σε περίπτωση πρόκλησης βλάβης. Το μεγάλο ζήτημα είναι η προστασία της ασφάλειας και των δεδομένων. Τα ρομπότ δεν μπορούν να λειτουργήσουν χωρίς ανταλλαγή δεδομένων, επομένως προκύπτει το ερώτημα για το ποιος θα έχει πρόσβαση σε αυτά τα δεδομένα.

Όσο αφορά την ευθύνη για τυχόν βλάβες που θα προκαλέσουν τα ρομπότ, είπε ότι δίνουν δύο επιλογές. Σύμφωνα με την αρχή της αντικειμενικής ευθύνης θα πρέπει να είναι ο κατασκευαστής ο υπεύθυνος, γιατί αυτός είναι υπεύθυνος για τον περιορισμό της πιθανότητας βλάβης. Μετά μπορεί να στραφεί εναντίον των παροχών του προϊόντος. Η άλλη επιλογή είναι η διεξαγωγή δοκιμών στα ρομπότ σε συνθήκες προσομοίωσης της πραγματικότητας για τον εντοπισμό και την αξιολόγηση των κινδύνων που αυτά συνιστούν.

Επίσης, αναφέρθηκε και στη συναισθηματική σύνδεση που μπορεί να αναπτυχθεί ανάμεσα σε εύαλωτα άτομα και σε ρομπότ. Πρέπει πάντα να λέμε στους ανθρώπους ότι τα ρομπότ δεν είναι αληθινά πρόσωπα, ότι μπορούν να δείξουν συμπάθεια αλλά δεν μπορούν να την νιώσουν. Δεν θέλουμε ρομπότ σαν αυτά που έχουν στην Ιαπωνία, που μοιάζουν με ανθρώπους. Προτείνουμε τη δημιουργία «χάρτη περί ρομποτικής», που θα θέτει τις βάσεις για τον εντοπισμό, την επισκόπηση και την τήρηση των θεμελιωδών δεοντολογικών αρχών ήδη από τη φάση του σχεδιασμού και της ανάπτυξης, συνέχισε. Πιστεύει επίσης, ότι η μεγαλύτερη πρόκληση για την κοινωνία μας και για το εκπαιδευτικό μας σύστημα, θα είναι οι θέσεις εργασίας, που στο μέλλον θα είναι λιγότερες και με υψηλή ειδικευση. Δεν ξέρουμε τι θα συμβεί είπε. Πρέπει να παρακολουθούμε την εξέλιξη και να είμαστε έτοιμοι για κάθε σενάριο.¹⁷⁰

Η πιθανότητα να ανατεθεί στις μηχανές η άδεια φόνου, έχει θορυβήσει κυβερνήσεις και ειδικούς που ασχολούνται με την ηθική, την ρομποτική και την ανθρώπινη ασφάλεια. Το 2009, κάποιος από αυτούς, συνέβαλαν στο να συγκροτηθεί η «Διεθνής Επιτροπή για τον Έλεγχο Ρομποτικών Όπλων (International Committee for Robot Arms Control)». Το 2012, οι Μη Κυβερνητικοί Οργανισμοί εξέφρασαν τον φόβο τους για τη χρήση και τους κινδύνους που προκαλούν τα αυτόματα όπλα. Αυτός ήταν και ο λόγος της σύνταξης μίας έκθεσης, πάνω σε αυτό, από το «Παρατηρητήριο Ανθρωπίνων Δικαιωμάτων (Human Rights Watch)». Στο Κοινοβούλιο του Λονδίνου, τον Απρίλιο του 2013, έκανε την εμφάνισή της η «Εκστρατεία για να Σταματήσουν τα Ρομπότ Φονιάδες (Campaign to Stop Killer Robots)».¹⁷¹ Με πρωτοστάτη την αντίθετη στην χρήση ναρκών, Jody Williams (κάτοχο του Νόμπελ Ειρήνης το 1997), η οργάνωση αυτή κατέληξε σε μια συνεργασία περισσότερων από τριάντα ΜΚΟ. Μετά από αυτό ο «Οργανισμός Ηνωμένων Εθνών» έκδωσε μια έκθεση¹⁷² που αφορούσε την επιβράδυνση της ανάπτυξης των «θανατηφόρων αυτόνομων ρομπότ».

Αυτή η εκστρατεία εναντίων των ρομπότ φονιάδων, είναι η μοναδική που εξεγείρεται ενάντια σε ένα όπλο το οποίο όχι μόνο δεν είναι πλήρως ανεπτυγμένο, αλλά ούτε η πρόκληση μαζικής ανθρωπιστικής βλάβης από αυτό, είναι αποδεδειγμένη. Η ανθρωπότητα όμως, φοβούμενη τα αποτελέσματα μιας τέτοιας τεχνολογίας, βασίζομενη στην ιδέα ότι «απλώς

¹⁷⁰ Ευρωπαϊκό Κοινοβούλιο, «Θέσπιση κανόνων για ρομπότ και τεχνητή νοημοσύνη ζητούν οι ευρωβουλευτές», Φεβρουάριος 2017

¹⁷¹ <https://www.stopkillerrobots.org>

¹⁷² <http://www.ohchr.org/Documents/HRBodies>

επειδή μπορούμε να κάνουμε κάτι, δεν σημαίνει ότι θα πρέπει και να το κάνουμε»¹⁷³, έβαλε τις βάσεις για ένα μελλοντικό έλεγχο των εξοπλισμών.



Εικόνα 22:
Αυτοκινούμενο όχημα.

Ένα επιπλέον μέτρο για την προστασία των ανθρώπων από τα υπολογιστικά συστήματα είναι και το νομοσχέδιο με το οποίο τίθενται οι όροι με τους οποίους επιτρέπεται η κυκλοφορία αυτο-οδηγούμενων αυτοκινήτων (self-driving cars) στην αγορά των ΗΠΑ. Η Βουλή των Αντιπροσώπων των ΗΠΑ ενέκρινε ομόφωνα το νομοθέτημα, με την ονομασία SELF DRIVE Act(link is external) (Safely Ensuring Lives Future Deployment and Research In Vehicle Evolution), το οποίο θα επιτρέψει την κυκλοφορία έως και 100.000 αυτόνομων οχημάτων στην αγορά των ΗΠΑ σε ετήσια βάση, εφόσον πληρούν τα ομοσπονδιακά κανονιστικά πρότυπα.

Σκοπός του νομοσχεδίου είναι να επιταχυνθεί ο έλεγχος των αυτοματοποιημένων συστημάτων και να καθοριστούν με ενιαίο τρόπο τα απαιτούμενα πρότυπα ασφαλείας, τα οποία είχαν αντιμετωπίσει διάφορα ζητήματα σε πολιτειακό επίπεδο. Στην πράξη το νομοσχέδιο εμποδίζει τις πολιτείες από το να μπλοκάρουν ή να επιβάλλουν πρόσθετες απαιτήσεις για τα αυτόνομα οχήματα, οι οποίες δεν είναι σύμφωνες με όσες ορίζονται στις επίσημες κατευθυντήριες γραμμές.

Το νομοσχέδιο προβλέπει επίσης τη δημιουργία ενός ειδικού ομοσπονδιακού συμβουλίου, ενώ τα δεδομένα που θα συλλέγονται από τα οχήματα θα αξιοποιούνται για τον καθορισμό καλύτερων κανονισμών και χαρακτηριστικών ασφαλείας. Μεταξύ άλλων, το νομοσχέδιο θεσπίζει υποχρέωση του κατασκευαστή για τη λήψη συγκεκριμένων μέτρων κυβερνοασφάλειας, αλλά και προστασίας της ιδιωτικής ζωής των ιδιοκτητών των αυτόνομων οχημάτων και των μεταφερόμενων προσώπων. Ωστόσο ορισμένες ενώσεις καταναλωτών εξέφρασαν έντονες ανησυχίες αναφορικά με τις δικλείδες ασφαλείας που θεσπίζονται.¹⁷⁴

¹⁷³ Council on Foreign Relations, Charli Carpenter, «Προσοχή στα ρομπότ δολοφόνους», Ιούλιο 2013

¹⁷⁴ Lawspot.gr, «Υπερψηφίστηκε ομόφωνα νομοσχέδιο για τα οχήματα χωρίς οδηγό από την βουλή των Αντιπροσώπων στις ΗΠΑ», Νοέμβριος 2017

7.5 Προτάσεις επιστημόνων για μελλοντικές ενέργειες.

Ο συνηθισμένος φόβος είναι, ότι οι έξυπνες μηχανές θα στραφούν ενάντια των ανθρώπων, λέει ο Hutan Ashrafian (Ο Hutan Ashrafian είναι κλινικός λέκτορας στη χειρουργική στο Imperial College London, UK, και Επιστημονικός Σύμβουλος στο Τμήμα Ιατρικής, Τμήμα Χειρουργικής & Καρκίνου). Αναρωτιέται όμως, ποιος μπορεί να σώσει τα ρομπότ από τα ίδια και από εμάς. Ισχυρίζεται ότι οι επιστήμονες, οι φιλόσοφοι, οι χρηματοδότες και οι σχεδιαστές πολιτικής, πρέπει να πάνε ένα στάδιο περαιτέρω και να εξετάσουν τις αλληλεπιδράσεις των ρομπότ με ένα άλλο ρομπότ και τις αλληλεπιδράσεις με την τεχνητή νοημοσύνη (AIonAI). Μαζί, πρέπει να αναπτύξουν μια πρόταση για έναν διεθνή χάρτη για τη τεχνητή νοημοσύνη, ισοδύναμο με εκείνον της Οικουμενικής Διακήρυξης των Ανθρώπινων Δικαιωμάτων των Ηνωμένων Εθνών. Ένα ανθρωπιστικό δίκαιο της τεχνητής νοημοσύνης και της ρομποτικής, θα μπορούσε να βοηθήσει να κατευθύνουν την έρευνα και την ανάπτυξη στην ηθικά διακριτική ρομποτική και εφαρμοσμένη μηχανική της τεχνητής νοημοσύνης.

Προτείνεται λοιπόν ένας νέος ρομποτικός νόμος, ο οποίος ονομάζεται «AIonAI» ή τεχνητή νοημοσύνη στην τεχνητή νοημοσύνη. Αυτός ο νόμος αντιμετωπίζει την αγνοημένη περιοχή όπου η μελλοντική τεχνητή νοημοσύνη είναι πιθανό να αλληλοεπιδράσει μεταξύ των ρομπότ, οδηγώντας ενδεχομένως στην εκμετάλλευση. Ως εκ τούτου, όλοι θα επωφεληθούν από την υιοθέτηση ενός καθολικού νόμου των δικαιωμάτων να αναγνωριστεί η έμφυτη αξιοπρέπεια και τα αναπαλλοτρίωτα δικαιώματα της Τεχνητής Νοημοσύνης. Μια τέτοια εξέταση μπορεί να βοηθήσει στο να αποτρέψει την εκμετάλλευση και την κακοποίηση των ορθολογικών και αισθανόμενων όντων, αλλά θα μπορούσε επίσης να απεικονίσει στον ηθικό μας κώδικα δεοντολογίας και την ανθρωπιά του πολιτισμού μας.¹⁷⁵

Οι εθνικές και διεθνείς τεχνολογικές πολιτικές πρέπει να εισαγάγουν τις έννοιες AIonAI στα τρέχοντα προγράμματα που στοχεύουν στην ανάπτυξη ασφαλών συστημάτων τεχνητής νοημοσύνης (AIs). Πρέπει να δεσμευτούμε με τις εκπαιδευτικές δραστηριότητες και την έρευνα, και να συνεχίσουμε να βελτιώνουμε τη φιλοσοφική συνείδηση. Θα μπορούσε ακόμη και να υπάρξει ένα ετήσιο βραβείο AIonAI για το «πιο αλτρουιστικά σχεδιασμένο AI». Οι κοινωνικοί επιστήμονες και οι φιλόσοφοι πρέπει να συνδεθούν με την τελευταία τεχνολογία ρομποτική και την έρευνα υπολογιστών. Οι τεχνολογικοί χρηματοδότες θα μπορούσαν να υποστηρίξουν ηθικές μελέτες για τις έννοιες AIonAI εκτός από τη χρηματοδότηση της ανάπτυξης της AI. Οι ιατρικοί χρηματοδότες όπως η «Wellcome Trust» ακολουθούν αυτό το πρότυπο ήδη, υποστηρίζοντας την έρευνα και για την υγειονομική περίθαλψη και για την ιατρική δεοντολογία και ιστορία.

Οι τρέχουσες και οι μελλοντικές ερευνητικές κοινότητες τεχνητής νοημοσύνης (AI) και ρομποτικής, πρέπει να στηρίξουν την έκθεση των ιδεών AIonAI. Η επόμενη γενιά των μηχανικών ρομποτικής και των ερευνητών της AI μπορεί επίσης να υιοθετήσει τις αρχές της AIonAI μέσω μίας σειράς μαθημάτων. Παραδείγματος χάριν, πολλοί άνθρωποι που θέλουν να ασχοληθούν με την βρετανική πολιτική, παίρνουν μια σειρά μαθημάτων την «PPE» (πολιτική, φιλοσοφία και οικονομικά). Μια ισοδύναμη σειρά μαθημάτων για τους σπουδαστές με τις φιλοδοξίες στη ρομποτική και το AI, θα μπορούσε να είναι η «CEP» (επιστήμη των υπολογιστών, εφαρμοσμένη μηχανική και φιλοσοφία), υποστηρίζει ο Hutan Ashrafian.

Θεωρώντας ότι θα έπρεπε να επεκτείνουμε τους τρεις νόμους της ρομποτικής του Asimov για να υποστηρίξουν το έργο σχετικά με την αλληλεπίδραση της AIonAI. Ο

¹⁷⁵ Hutan Ashrafian, *Science and Engineering Ethics «AIonAI: A Humanitarian Law of Artificial Intelligence and Robotics»*, 21 (1):29-40 (2015)

συγγραφέας και λέκτορας, προτείνει έναν ακόμα νόμο. Ισχυρίζεται ότι: «όλα τα ρομπότ που είναι προικισμένα με συγκρίσιμη ανθρώπινη λογική και συνείδηση, θα πρέπει να συμπεριφέρονται μεταξύ τους με πνεύμα αδελφοσύνης και αδελφότητας. Μην υποτιμάτε την πιθανότητα της τεχνητής σκέψης των μηχανών. Η ανθρωπότητα φθάνει στον ορίζοντα της γέννησης μιας νέας ευφυούς φυλής. Το εάν αυτή η νοημοσύνη είναι ή όχι «τεχνητή» δεν αναιρεί το ζήτημα ότι ο νέος ψηφιακός λαός θα αξίζει την ηθική αξιοπρέπεια και τα δικαιώματα, και έναν νέο νόμο για να τους προστατεύσει».¹⁷⁶

Τις ίδιες ανησυχίες έχουν και οι ευρωβουλευτές. Στην έκθεσή τους προς την επιτροπή του Ευρωκοινοβουλίου, της 31/05/2016, σημειώνουν ότι το αναπτυξιακό δυναμικό μέσω της χρήσης της ρομποτικής χαρακτηρίζεται από εντάσεις και κινδύνους και θα πρέπει να εκτιμηθεί προσεκτικά από πλευράς ανθρώπινης ασφάλειας, υγείας, ελευθερίας, ιδιωτικής ζωής, ακεραιότητας και αξιοπρέπειας, αυτονομίας και μη εισαγωγής διακρίσεων καθώς και από πλευράς προστασίας των προσωπικών δεδομένων.

Φρονεί ότι το ισχύον ρυθμιστικό πλαίσιο της Ένωσης πρέπει να εκσυγχρονιστεί και να συμπληρωθεί, όπου αυτό είναι απαραίτητο, βάσει ηθικών αρχών που θα αποτυπώνουν τον πολύπλοκο χαρακτήρα του τομέα της ρομποτικής και των πολυάριθμων επιπτώσεών της στην κοινωνία, την ιατρική και τη βιοηθική. Ότι είναι απαραίτητο ένα σαφές, αυστηρό και αποτελεσματικό δεοντολογικό πλαίσιο με κατευθυντήριες γραμμές για την εξέλιξη, τον σχεδιασμό, την παραγωγή, τη χρήση και την τροποποίηση των ρομπότ. Πρότείνει λοιπόν, να περιληφθεί στο παράρτημα του ψηφίσματος ένα πλαίσιο υπό μορφή χάρτη, αποτελούμενο από έναν κώδικα δεοντολογίας για μηχανικούς ρομποτικής, έναν κώδικα για τις επιτροπές δεοντολογίας σε θέματα έρευνας για την επανεξέταση των πρωτοκόλλων ρομποτικής, καθώς και από υποδείγματα αδειών για σχεδιαστές και χρήστες.

Υπογραμμίζει την αρχή της διαφάνειας, δηλαδή ότι θα πρέπει πάντα να είναι δυνατό να παρουσιαστεί το σκεπτικό πίσω από κάθε απόφαση που λαμβάνεται με τη βοήθεια της τεχνητής νοημοσύνης, η οποία μπορεί να έχει σημαντικό αντίκτυπο στη ζωή ενός ή περισσότερων ανθρώπων. Πάντα πρέπει να προσφέρεται η δυνατότητα παρουσίασης των υπολογισμών του συστήματος τεχνητής νοημοσύνης σε μορφή κατανοητή από τον άνθρωπο. Φρονεί ότι τα προηγμένα ρομπότ πρέπει να είναι εξοπλισμένα με ένα «μαύρο κουτί», το οποίο θα καταγράφει στοιχεία για κάθε λειτουργία που εκτελείται από το ρομπότ, συμπεριλαμβανομένων των λογικών σταδίων που συνέβαλαν στη λήψη των αποφάσεών του.

Οι κατευθυντήριες γραμμές του δεοντολογικού πλαισίου πρέπει να βασίζονται στις αρχές της ωφέλιμης δράσης, της αποτροπής ζημιών, της αυτονομίας και της δικαιοσύνης καθώς και στις υφιστάμενες δεοντολογικές πρακτικές και τους σχετικούς κώδικες. Επίσης, στις αρχές και αξίες που κατοχυρώνονται στο άρθρο 2 της Συνθήκης για τη λειτουργία της Ευρωπαϊκής Ένωσης και στον Χάρτη των Θεμελιωδών Δικαιωμάτων της ΕΕ, όπως είναι η ανθρώπινη αξιοπρέπεια, η ισότητα, η δικαιοσύνη και η επιείκεια, η απαγόρευση των διακρίσεων, η συνειδητή συγκατάθεση, η προστασία της ιδιωτικής ζωής και των δεδομένων καθώς και σε άλλες βασικές αρχές και αξίες του γνωστικού δικαίου, όπως η απαγόρευση στιγματισμού, η διαφάνεια, η αυτονομία, η ατομική και κοινωνική ευθύνη.

Ο Hutan προτείνει στην Επιτροπή να εξετάσει τη σύσταση ενός Ευρωπαϊκού Οργανισμού για τη ρομποτική και την τεχνητή νοημοσύνη, ο οποίος θα παρέχει την απαραίτητη τεχνογνωσία και εμπειρογνωμοσύνη σε ρυθμιστικά και δεοντολογικά θέματα για τη στήριξη των οικείων δημόσιων φορέων, τόσο σε επίπεδο Ένωσης όσο και σε επίπεδο κρατών μελών. Αυτό θα βοηθήσει στο να ενισχυθούν οι προσπάθειές τους να διασφαλίσουν μια έγκαιρη, δεοντολογική και κατόπιν δέουσα ενημέρωση αντίδραση στις νέες ευκαιρίες

¹⁷⁶ Hutan Ashrafian, nature.com, « Intelligent robots must uphold human rights», Nature 519, 391, doi:10.1038/519391a, Μάρτιος 2015

και προκλήσεις, διασυνοριακού ιδίως χαρακτήρα, που προκύπτουν από τις τεχνολογικές εξελίξεις στον τομέα της ρομποτικής, όπως ο τομέας των μεταφορών. Επίσης προτείνει να δοθεί σε αυτόν τον Οργανισμό ο κατάλληλος εξοπλισμός και να στελεχωθεί με εξωτερικούς εμπειρογνώμονες σε τεχνικά και δεοντολογικά ζητήματα, ώστε να ασχολούνται με τον διατομεακό και διεπιστημονικό έλεγχο των εφαρμογών που βασίζονται στη ρομποτική, την αναζήτηση προτύπων βέλτιστων πρακτικών, και, κατά περίπτωση, να προτείνουν ρυθμιστικά μέτρα, να ορίζουν νέες βασικές αρχές και να αντιμετωπίζουν δυνητικά προβλήματα προστασίας των καταναλωτών και συστηματικές προκλήσεις. Προτείνει ακόμα, να υποβάλλουν ετησίως έκθεση προς το Ευρωπαϊκό Κοινοβούλιο σχετικά με τις τελευταίες εξελίξεις στον τομέα της ρομποτικής, καθώς και με τα πιθανά μέτρα που πρέπει να ληφθούν. Επίσης, καλεί την Επιτροπή και τα κράτη μέλη να αναπτύξουν κατευθυντήριες γραμμές για να βοηθήσουν στη σύσταση και τη λειτουργία των εν λόγω επιτροπών.

Όσον αφορά τον τομέα των ιατρικών εφαρμογών ζωτικής σημασίας, όπως των ρομποτικών τεχνητών μελών, επισημαίνει τον κίνδυνο που συνδέεται με το ενδεχόμενο παράνομης πρόσβασης, διακοπής λειτουργίας ή διαγραφής δεδομένων σε κυβερνοφυσικά συστήματα (CPS) εμφυτευμένα στο ανθρώπινο σώμα, καθώς κάτι τέτοιο θα μπορούσε να θέσει σε κίνδυνο την υγεία ή, σε ακραίες περιπτώσεις, ακόμα και τη ζωή των ανθρώπων και τονίζει ως εκ τούτου ότι πρέπει να δοθεί προτεραιότητα στην προστασία τέτοιων συστημάτων. Συνιστά να δημιουργηθούν ανεξάρτητες καταπιστευτικές οντότητες που θα διαθέτουν τα απαραίτητα μέσα για να παρέχουν υπηρεσίες φροντίδας, όπως συντήρηση, επισκευές και βελτιώσεις, συμπεριλαμβανομένων αναβαθμίσεων λογισμικού σε άτομα που φέρουν ζωτικές και προηγμένες ιατρικές συσκευές, ειδικά όταν οι εν λόγω υπηρεσίες δεν παρέχονται πλέον από τον αρχικό προμηθευτή. Τονίζει ότι είναι σημαντικό να διασφαλίζεται η ισότιμη πρόσβαση για όλους τους ανθρώπους στις εν λόγω τεχνολογικές καινοτομίες, τα μέσα και τις επεμβάσεις και καλεί την Επιτροπή και τα κράτη μέλη να προωθήσουν την ανάπτυξη υποστηρικτικών τεχνολογιών, προκειμένου να προωθηθεί η ανάπτυξη και η χρήση των εν λόγω τεχνολογιών από άτομα που την έχουν ανάγκη, σύμφωνα με το άρθρο 4 της σύμβασης του ΟΗΕ για τα δικαιώματα των ατόμων με αναπηρία, στην οποία η Ένωση είναι συμβαλλόμενο μέρος.

Προτείνει στα κράτη μέλη να αναπτύξουν πιο ευέλικτα συστήματα κατάρτισης και εκπαίδευσης, έτσι ώστε οι στρατηγικές δεξιότητες να ανταποκρίνονται στις ανάγκες της αναπτυσσόμενης ρομποτικής οικονομίας. Επισημαίνει τις προβλέψεις της Επιτροπής, σύμφωνα με τις οποίες ως το 2020 η Ευρώπη μπορεί να αντιμετωπίζει έλλειψη έως και 825.000 επαγγελματιών στον τομέα της τεχνολογίας των πληροφοριών και επικοινωνιών, ενώ στο 90% των θέσεων εργασίας θα απαιτούνται τουλάχιστον βασικές ψηφιακές δεξιότητες. Καλεί την Επιτροπή να παράσχει ουσιαστική υποστήριξη για την ανάπτυξη ψηφιακών δεξιοτήτων σε όλες τις ηλικιακές ομάδες και ανεξαρτήτως εργασιακού καθεστώτος ως ένα πρώτο βήμα για την καλύτερη εξισορρόπηση μεταξύ ελλείψεων στην αγορά εργασίας. Ζητάει από την Επιτροπή και τα κράτη μέλη να αναλάβουν πρωτοβουλίες ώστε να στηρίξουν την παρουσία των γυναικών στον τομέα της τεχνολογίας των πληροφοριών και επικοινωνιών και να δώσουν ώθηση στις ηλεκτρονικές τους δεξιότητες. Τονίζει τη σημασία που έχει η ευελιξία των δεξιοτήτων καθώς και οι κοινωνικές, δημιουργικές και ψηφιακές δεξιότητες στην εκπαίδευση, δεδομένου ότι, εκτός από την ακαδημαϊκή γνώση που παρέχεται στα πανεπιστήμια, η δια βίου μάθηση πρέπει να επιτυγχάνεται και με μια δια βίου δραστηριότητα.

Η επιτροπή θα πρέπει να αρχίσει να αναλύει και να παρακολουθεί στενότερα τις μεσοπρόθεσμες και μακροπρόθεσμες τάσεις στον τομέα της απασχόλησης, εστιάζοντας στη δημιουργία, τη μεταφορά και την απώλεια θέσεων εργασίας σε διάφορους τομείς δεξιοτήτων, έτσι ώστε να διαπιστωθούν οι τομείς, στους οποίους δημιουργούνται και οι τομείς στους οποίους καταργούνται θέσεις εργασίας, ως συνέπεια της αυξημένης χρήσης ρομπότ. Να αναλύσει μια σειρά δυνητικών εξελίξεων και τις συνέπειές τους, αλλά και να ξεκινήσει ένας

διάλογος χωρίς αποκλεισμούς όσον αφορά τα νέα πρότυπα απασχόλησης και τη βιωσιμότητα των φορολογικών συστημάτων και των συστημάτων κοινωνικής ασφάλισης.

Θεωρεί ότι η αστική ευθύνη για ζημιές που προκαλούνται από τα ρομπότ αποτελεί ζήτημα ζωτικής σημασίας που πρέπει να εξεταστεί και να ρυθμιστεί σε επίπεδο Ένωσης, έτσι ώστε να διασφαλιστεί ο ίδιος βαθμός αποτελεσματικότητας, διαφάνειας συνοχής και δικαιοσύνης, σε ολόκληρη την Ευρωπαϊκή Ένωση, προς όφελος των πολιτών, των καταναλωτών και των επιχειρήσεων. Σημειώνει ότι η κοινή δράση ανθρώπου-ρομπότ, πρέπει να διέπεται από δύο βασικές σχέσεις αλληλεξάρτησης, την προβλεψιμότητα και την καθοδηγησιμότητα και επισημαίνει ότι αυτές οι δύο σχέσεις αλληλεξάρτησης έχουν καίρια σημασία για τον καθορισμό των πληροφοριών που πρέπει να ανταλλάσσουν άνθρωποι και ρομπότ, καθώς και για τον προσδιορισμό μιας κοινής βάσης μεταξύ ανθρώπων και ρομπότ προκειμένου να καταστεί εφικτή η ομαλή κοινή τους δράση.

Το μελλοντικό νομοθετικό μέσο πρέπει να βασιστεί σε ενδεδειγμένη αξιολόγηση από την Επιτροπή, που θα καθορίσει εάν πρέπει να εφαρμοστεί μια προσέγγιση με βάση την αντικειμενική ευθύνη ή τη διαχείριση κινδύνου. Ταυτόχρονα η αντικειμενική ευθύνη απαιτεί απόδειξη μόνο για την ύπαρξη ζημίας και την αιτιώδη συνάφεια μεταξύ της ζημιογόνου λειτουργίας του ρομπότ και της ζημίας που υπέστη ο παθών. Ενώ, μια προσέγγιση με βάση τη διαχείριση κινδύνου δεν εστιάζει στο πρόσωπο «που ενήργησε αμελώς» ως ατομικά υπεύθυνο, αλλά στο πρόσωπο που μπορεί, υπό ορισμένες συνθήκες, να ελαχιστοποιήσει τους κινδύνους και να αντιμετωπίσει τις αρνητικές συνέπειες.

Καλεί τον ασφαλιστικό τομέα να αναπτύξει νέα προϊόντα και προσφορές, που λαμβάνουν υπόψη την πρόοδο της ρομποτικής. Προτείνει τη δημιουργία υποχρεωτικού συστήματος ασφάλισης, όπου είναι σκόπιμο και απαραίτητο για συγκεκριμένες κατηγορίες ρομπότ, παρόμοιου με το υφιστάμενο σύστημα ασφάλισης των αυτοκινήτων, στο πλαίσιο του οποίου θα απαιτείται από τους κατασκευαστές και τους ιδιοκτήτες ρομπότ να συνάπτουν σύμβαση ασφαλιστικής κάλυψης για ζημιές που ενδέχεται να προκαλέσουν τα ρομπότ τους.

Τέλος, υπογραμμίζει τη σημασία της άμεσης σύστασης σε νοσοκομεία και άλλα ιδρύματα υγειονομικής περίθαλψης, κατάλληλα στελεχωμένων επιτροπών για θέματα ηθικής στον τομέα της ρομποτικής, οι οποίες θα είναι επιφορτισμένες με το καθήκον να εξετάζουν και να συμβάλλουν στην επίλυση ασυνήθιστων, περίπλοκων ηθικών ζητημάτων που επηρεάζουν τη φροντίδα και τη θεραπεία των ασθενών. Προκειμένου να διαχειριστούμε αυτή την πραγματικότητα και να διασφαλίσουμε ότι τα ρομπότ είναι και θα παραμείνουν στην υπηρεσία του ανθρώπου, οι ευρωπαίοι νομοθέτες υποστήριξαν ότι ήρθε η ώρα να ξεκινήσουν οι διαδικασίες για την θέσπιση κανόνων αναφορικά με τη χρήση των υπολογιστικών συστημάτων.¹⁷⁷

Αυτό που χρειάζεται είναι «ένα πλαίσιο ανάπτυξης, βασικούς κανόνες και γενικές προδιαγραφές που προσδίδουν διαφάνεια, αξιοπιστία και ασφάλεια». Την ίδια στιγμή, είναι απαραίτητο ο κόσμος να κατανοήσει την τεχνολογία, τις αλλαγές που θα επέλθουν στην κοινωνία, τις ευκαιρίες που θα δημιουργήσει, καθώς και τους κινδύνους που θα προκαλέσει. Ο κρατικός φορέας πρέπει να ενεργοποιήσει κατάλληλους κρατικούς μηχανισμούς, δημιουργώντας νέες στρατηγικές, τροποποιώντας το εκπαιδευτικό σύστημα, καθώς και την λειτουργία των αγορών. Τέλος, απαραίτητη είναι η δημιουργία ενός δυνατού πλαισίου που θα συμβάλλει στην «αποφυγή της δυσανάλογης συγκέντρωσης τεχνολογικής δύναμης και ελέγχου».

¹⁷⁷ ΣΧΕΔΙΟ ΕΚΘΕΣΗΣ που περιέχει συστάσεις προς την Επιτροπή για ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής (2015/2013(INL)) Επιτροπή Νομικών Θεμάτων Εισηγήτρια: Mady Delvaux (Πρωτοβουλία – άρθρο 46 του Κανονισμού), Μάιος 2016



Εικόνα 23:

Διανύουμε την εποχή που οι άνθρωποι πρέπει να εξετάσουν προσεκτικά πολλαπλές λεπτομέρειες πριν μπορέσουν να κρίνουν αν ένα άτομο ή ένα αντικείμενο είναι ζωντανό.

Ακολουθούν τα Συμπεράσματα και η Βιβλιογραφία.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Μέρα με την μέρα γίνεται όλο και πιο σαφές ότι δεν πρέπει μόνο οι άνθρωποι να έχουν ηθικά δικαιώματα αλλά και τα άλλα πλάσματα, τεχνητά ή φυσικά που ζουν στον πλανήτη Γη. Τα ανθρώπινα όντα προσπαθούν να καθορίσουν ποια είναι τα ηθικά προβλήματα και πώς μπορούν να κατασκευάσουν μια τεχνητή νοημοσύνη, ώστε να παίρνει σωστές αποφάσεις, όχι μόνο για ένα άτομο αλλά και για την κοινωνία ολόκληρη. Επίσης γίνεται όλο και πιο σημαντική η αναγνώριση της σημασίας, του ρόλου και της θέσης της τεχνητής νοημοσύνης στην κοινωνία, κάτι το οποίο δεν αποτελεί υποκατάστατο για την ανθρώπινη νοημοσύνη.

Όσο προσεγγίζεται χρονικά η «μοναδικότητα» τόσο απαραίτητη καθίσταται η υιοθέτηση νομοθεσίας που θα ρυθμίζει την συμπεριφορά των τεχνητών νοήμων όντων. Μέχρις ότου η ηθική γίνει ζωτικό μέρος της ανθρώπινης συμπεριφοράς, δεν υπάρχει κάποια απάντηση σχετικά με την ασφάλεια μίας υπολογιστικής νοημοσύνης. Η υπολογιστική νοημοσύνη και τα ρομποτικά συστήματα θεωρείται ότι έχουν την ικανότητα να ικανοποιήσουν οποιαδήποτε ανθρώπινη απαίτηση. Στις περιπτώσεις μάλιστα που έχουν και ανθρώπινα χαρακτηριστικά, αυτό επηρεάζει τους ανθρώπους με τέτοιον τρόπο, ώστε να τους κάνει να θεωρήσουν ότι θα έχουν επίσης ανθρώπινες τάσεις, συναισθήματα και τρόπο σκέψης. Αυτό από μόνο του δημιουργεί καταστάσεις που δεν έχει βιώσει η ανθρωπότητα μέχρι τώρα και έτσι προκύπτουν τα ηθικά, νομικά και κοινωνικά προβλήματα.

Αυτές οι καταστάσεις, θα δημιουργήσουν μικρές ή μεγάλες επιπτώσεις στην ανθρώπινη καθημερινότητα, που θα πρέπει να εξεταστούν και να δοθούν λύσεις, όταν τα ρομπότ αντικαταστήσουν την ανθρώπινη φροντίδα και συντροφιά. Πρέπει να δοθεί ιδιαίτερη προσοχή στο ενδεχόμενο να αναπτυχθεί μια εξαρτημένη συναισθηματική σχέση μεταξύ ανθρώπων και ρομπότ, ιδιαίτερα σε ευπαθείς ομάδες, όπως παιδιά, ηλικιωμένους και άτομα με αναπηρία. Η προσκόλληση αυτή είναι πιθανόν να αυξήσει τις πιθανότητες δημιουργίας σοβαρών συναισθηματικών ή σωματικών προβλημάτων.

Η ανθρωπότητα βρίσκεται πλέον στο κατώφλι μιας εποχής, κατά την οποία οι ραγδαίες εξελίξεις στον τομέα της υπολογιστικής νοημοσύνης, αναμένεται να πυροδοτήσουν μια νέα βιομηχανική επανάσταση, η οποία είναι πιθανόν να μην αφήσει ανέγγιχτο κανένα κοινωνικό στρώμα. Η εξέλιξη των ρομπότ και της τεχνητής νοημοσύνης μπορεί να προκαλέσει, σε μεγάλο βαθμό, την αντικατάσταση της ανθρώπινης εργασίας, πράγμα που δημιουργεί ανησυχίες σχετικά με το μέλλον της απασχόλησης και της επιβίωσης των συστημάτων κοινωνικής ασφάλισης, αλλά και της διεύρυνσης των ανισοτήτων στην κατανομή του πλούτου.

Οι μεγάλες τεχνολογικές εταιρίες, έχουν επενδύσει πολλά στην επιστημονική έρευνα και στην εμπορική ανάπτυξη της τεχνητής νοημοσύνης. Ισχυρό τους κίνητρο για την επικράτηση στην αγορά, είναι η εξαγορά ανερχόμενων, μικρών, τεχνολογικών εταιριών με σκοπό την δημιουργία τεχνολογικών υπερδυνάμεων, αποσκοπώντας στην συγκέντρωση πλούτου και δύναμης.

Χαρακτηριστικά είναι τα λόγια που είπε πρόσφατα ο επιχειρηματίας Έλον Μασκ, ιδρυτής της Tesla και της Space X, που μιλώντας στο πλαίσιο ενός νέου του ντοκιμαντέρ με τίτλο «Εμπιστευέστε αυτόν τον υπολογιστή;», εξέφρασε τις ανησυχίες του για την τεχνητή νοημοσύνη: «Έχοντας άμεση πρόσβαση στην παραμικρή πληροφορία για κάθε άνθρωπο του πλανήτη, μπορώντας να κάνει εκατομμύρια υπολογισμούς σε κλάσμα του δευτερολέπτου και να παίρνει κρίσιμες αποφάσεις, η τεχνητή νοημοσύνη ίσως κάποτε είναι σε θέση όχι μόνο να ελέγχει την παγκόσμια οικονομία και τα οπλικά συστήματα, αλλά τελικά να ενδυθεί τον μανδύα

του δικτάτορα. Θα υπάρξει ένας αθάνατος δικτάτορας από τον οποίο δεν θα μπορούμε ποτέ να ξεφύγουμε. Αν η τεχνητή νοημοσύνη έχει ένα στόχο και η ανθρωπότητα απλώς συμβαίνει να παρεμβάλλεται σε αυτόν, θα καταστρέψει την ανθρωπότητα σαν κάτι φυσικό, χωρίς καν να το σκεφτεί. Είναι σαν να φτιάχνουμε ένα δρόμο και ένα μυρμήγκι συμβεί να βρεθεί στο δρόμο μας. Δεν μισούμε τα μυρμήγκια, απλώς φτιάχνουμε ένα δρόμο. Κι, έτσι, αντίο μυρμήγκι...».¹⁷⁸

Επίσης, πολύπλοκα ερωτήματα εγείρουν και οι τεράστιες πρόοδοι στον τομέα της επισκευής και της αποκατάστασης οργάνων που έχουν υποστεί βλάβη, αλλά και άλλων ανθρωπίνων λειτουργιών, ιδίως λόγω των δυνατοτήτων για βελτιστοποίηση του ανθρώπου. Τα ιατρικά ρομπότ μπορούν να εμφυτευτούν στο ανθρώπινο σώμα και θα μπορούσαν να αλλάξουν ουσιαστικά την αντίληψη περί υγιούς ανθρώπινου σώματος. Εκείνο που δεν πρέπει να διαφύγει είναι η πιθανή απάλειψη των κληρονομικών διαφοροποιήσεων που θα χαθούν με τον καιρό, λόγω της τεχνολογικής επέμβασης του ανθρώπου.

Προσπαθώντας ο άνθρωπος να βελτιστοποιηθεί, επεμβαίνει τεχνολογικά στο σώμα του και ρομποτοποιείται. Με τον καιρό, όλοι οι άνθρωποι θα έχουν τις ίδιες σωματικές δυνατότητες και περίπου τον ίδιο χαρακτήρα αφού θα μπορούν να επέμβουν στον εγκέφαλο και να αποβάλλουν οτιδήποτε αρνητικό. Αν αυτό συμβεί, τότε δεν θα προκύψει θέμα μόνο για την τροποποίηση ενός ανθρώπου, αλλά για την τροποποίηση ολόκληρου του ανθρώπινου είδους.

Από την άλλη, έχει αποδειχτεί ότι ο τομέας της νοημοσύνης των μηχανών προσφέρει τεράστια οικονομικά και τεχνολογικά οφέλη στην κοινωνία, βελτιώνοντας σε πολύ μεγάλο βαθμό την ικανότητα ανάλυσης δεδομένων και της απλούστευσης των διαδικασιών λήψης αποφάσεων. Αυτό με τον καιρό, θα οδηγήσει στην ανάπτυξη έξυπνων και αυτόνομων μηχανών, οι οποίες θα έχουν την ικανότητα να τυγχάνουν εκπαίδευσης και να λαμβάνουν αποφάσεις μόνες τους, χωρίς εντολή. Αυτό το γεγονός όμως, δεν αποφέρει μόνο οικονομικά οφέλη, αλλά εγείρει και διάφορους προβληματισμούς σχετικά με τις άμεσες και έμμεσες επιπτώσεις στην κοινωνία ως σύνολο.

Η ανάπτυξη της ρομποτικής και της υπολογιστικής νοημοσύνης ενδέχεται να επιφέρουν μεταβολές στον τρόπο ζωής και εργασίας, να αυξήσουν την αποδοτικότητα και να βελτιώσουν την παροχή υπηρεσιών, όχι μόνο στους τομείς της παραγωγής και του εμπορίου, αλλά και σε τομείς όπως οι μεταφορές, το σύστημα υγείας, η εκπαίδευση, η κτηνοτροφία και η γεωργία. Πρέπει να αξιολογηθούν οι οικονομικές μεταβολές, όπως και ο αντίκτυπος της ρομποτικής και του τομέα της νοημοσύνης των μηχανών στην απασχόληση.

Παρά τα αναμφισβήτητα πλεονεκτήματα που προσφέρει η ρομποτική, οι εφαρμογές της ενδέχεται να επιφέρουν μεταβολές στην αγορά εργασίας, έχοντας σαν πιθανή συνέπεια, αλλαγές στα δεδομένα της εκπαίδευσης, της απασχόλησης και της κοινωνικής πολιτικής. Κάποιες από αυτές θα είναι: η αναδιάρθρωση κλάδων, ο αφανισμός κάποιων επαγγελμάτων, ο επαναπροσδιορισμός ωραρίων εργασίας και μοντέλων απασχόλησης. Στον αντίποδα, θα δημιουργηθούν νέα επαγγέλματα, ανάγκη για εξειδίκευση, δημιουργία νέων επιστημονικών πεδίων και δημιουργικότερη εκμετάλλευση του χρόνου.

Βέβαια, όπως εκτιμάται από κάποιους επιστήμονες, ο αυτοματισμός (μηχανοποίηση, ρομποτική, πληροφορική), δημιούργησε πολύ περισσότερα νέα επαγγέλματα από όσα εξαφάνισε. Απόδειξη είναι το γεγονός ότι τους τελευταίους δύο αιώνες, τα ποσοστά απασχόλησης αυξάνονται σταθερά λόγω των τεχνολογικών εξελίξεων. Έτσι, τα ρομπότ πιθανόν να δημιουργήσουν περισσότερες θέσεις εργασίας σε σύγκριση με αυτές που θα καταργήσει η χρήση τους, ωστόσο οι θέσεις που θα δημιουργηθούν θα είναι υψηλής ειδίκευσης

¹⁷⁸ ΑΜΠΕ (Αθηναϊκό Μακεδονικό Πρακτορείο Ειδήσεων), «Ο Μασκ προειδοποιεί πως η τεχνητή νοημοσύνη μπορεί να γίνει αθάνατος δικτάτορας», Απρίλιος 2018

Κανένας δεν μπορεί να παραβλέψει, βέβαια, το τεράστιο δυναμικό της υπολογιστικής νοημοσύνης για τη βελτίωση της ασφάλειας στον τόπο εργασίας. Με την χρήση των ρομπότ, επιτυγχάνεται η πραγματοποίηση πολλών επικίνδυνων και επιβλαβών εργασιών για την υγεία των ανθρώπων, όπως για παράδειγμα ο καθαρισμός χώρων μολυσμένων με τοξικά απόβλητα. Στο πλαίσιο όμως της ανάπτυξης της υπολογιστικής νοημοσύνης, είναι ζωτικής σημασίας να διασφαλιστεί ότι θα είναι ανά πάσα στιγμή δυνατή η άσκηση ανθρώπινου ελέγχου επί των έξυπνων μηχανών.

Είναι λοιπόν αναγκαίο να διερευνηθεί, κατά πόσο η ραγδαία εξέλιξη των υπολογιστικών συστημάτων συμβαδίζει με τους ηθικούς κανόνες την ανθρώπινης κοινωνίας. Θεωρείται πλέον δεδομένο ότι οι εφαρμογές και οι συσκευές θα επικοινωνούν μεταξύ τους, αλλά και με βάσεις δεδομένων χωρίς ανθρώπινη παρέμβαση, ή ακόμη και χωρίς να γίνεται καν αντιληπτό ότι επικοινωνούν. Αυτό καθιστά απαραίτητη και την εξέταση βασικών ζητημάτων που συνδέονται με την προστασία των δεδομένων, της ιδιωτικότητας και του ηλεκτρονικού εμπορίου.

Οι άνθρωποι, κατά τη διάρκεια της συνεργασίας τους με τα ρομπότ, δεν θα πρέπει να εκτίθενται σε επιπλέον κινδύνους σε σχέση με τους κινδύνους στους οποίους εκτίθενται στην καθημερινότητά τους. Η λειτουργία ενός συστήματος ρομποτικής πρέπει πάντα να βασίζεται σε διαδικασίες διεξοδικής αξιολόγησης κινδύνου. Οι μηχανικοί, κατασκευαστές και προγραμματιστές, έχουν μεγάλη ευθύνη. Πρέπει να λαμβάνουν υπόψη και να σέβονται τη σωματική ακεραιότητα, την ασφάλεια, την υγεία και τα δικαιώματα όλων των ανθρώπων. Πρέπει να διαφυλάσσονται από αυτούς, η ευημερία του ανθρώπου, ο σεβασμός των ανθρώπινων δικαιωμάτων και να δημοσιοποιούνται αμέσως παράγοντες που μπορεί να αποτελούν κίνδυνο για το κοινό ή για το περιβάλλον. Επίσης, να μην αποκλίνουν από τον αρχικό σκοπό ανάπτυξης της ρομποτικής τεχνολογίας, ο οποίος εστιάζεται στη συμπλήρωση των ανθρώπινων ικανοτήτων κι όχι στην αντικατάστασή τους.

Συνεπώς, οι σχεδιαστές ρομποτικής έχουν ευθύνη για την ανάπτυξη διαδικασιών τέτοιων ώστε, να διασφαλίζεται ότι δεν θα αποκαλύπτονται πληροφορίες για την ταυτότητα συγκεκριμένων ατόμων κι ότι οι ιδιωτικές πληροφορίες θα είναι ασφαλείς και θα χρησιμοποιούνται αποκλειστικά με τον κατάλληλο τρόπο. Το δικαίωμα στην ιδιωτική ζωή είναι ηθικό δικαίωμα και πρέπει να γίνεται σεβαστό, ακόμα κι από την υπολογιστική νοημοσύνη.

Ωστόσο, ποιες άραγε θα είναι οι συνέπειες από τη βλάβη ενός συστήματος ή την παραβίαση (hacking) συνδεδεμένων συστημάτων ρομποτικής ή αν ο κωδικός ενός ρομπότ αποδειχθεί λανθασμένος; Το υφιστάμενο νομικό πλαίσιο δεν επαρκεί για την κάλυψη της ζημιάς που προκαλείται από εξελιγμένα, αυτόνομα ρομπότ, εξοπλισμένα με ικανότητες προσαρμογής και μάθησης από τη δική τους εμπειρία και που μπορούν να έρχονται σε διάδραση με το περιβάλλον τους με απρόβλεπτο τρόπο. Συνεπώς, είναι εξαιρετικά σημαντικό να θεσπιστούν καινούριοι νόμοι, κατά τους οποίους ο νομοθέτης θα λάβει υπόψιν του τις νομικές και ηθικές συνέπειες, αλλά και τις επιπτώσεις της εξέλιξης αυτής, χωρίς να καταστέλλει βέβαια την καινοτομία.

Σε αυτό το σημείο πρέπει να αναφερθεί ότι στις μέρες μας σχεδιάζονται και κατασκευάζονται «έξυπνα ρομπότ» όχι μόνο για τη βιομηχανία, τις υπηρεσίες, και τις οικιακές εργασίες, όπως έχει προαναφερθεί, αλλά δυστυχώς και για πολεμικούς σκοπούς. Όπως και άλλα επιτεύγματα της επιστήμης που έχουν χρησιμοποιηθεί εναντίον του ανθρώπου, έτσι και τα ευφυή συστήματα έχουν χρησιμοποιηθεί και μπορούν να ξαναχρησιμοποιηθούν για καταστροφικούς σκοπούς. Για αυτόν τον λόγο, είναι αναγκαία η δημιουργία ενός ενιαίου, παγκόσμιου κώδικα δεοντολογικής συμπεριφοράς, προκειμένου να ρυθμιστούν θέματα ηθικής αλλά και ευθύνης, όσον αφορά τη στρατιωτική χρήση έξυπνων ρομποτικών συστημάτων, όπως επίσης θέματα που αφορούν τις κοινωνικές και περιβαλλοντικές επιπτώσεις της ρομποτικής στην ανθρώπινη υγεία και να διασφαλιστεί η

συμμόρφωσή της με νομικά και ηθικά πρότυπα. Όπως είπε ο Έλον Μασκ στο προαναφερόμενο ντοκιμαντέρ: *"Κατευθυνόμαστε ταχέως προς μια ψηφιακή υπερνοημοσύνη που θα υπερβαίνει κατά πολύ οποιονδήποτε άνθρωπο"* και προειδοποίησε ότι *"αν μια εταιρεία ή μια μικρή ομάδα ανθρώπων καταφέρει να αναπτύξει θεική ψηφιακή υπερνοημοσύνη, θα κυριαρχήσει στον κόσμο"*. Ανέφερε επίσης ότι: *"Αν η τεχνητή νοημοσύνη έχει ένα στόχο και η ανθρωπότητα απλώς συμβαίνει να παρεμβάλλεται σε αυτόν, θα καταστρέψει την ανθρωπότητα σαν κάτι φυσικό, χωρίς καν να το σκεφτεί"*. Φαίνεται πως το μέλλον είναι ήδη εδώ. Οι επιστήμονες πρέπει σύντομα να επιληφθούν του θέματος και να αποδεχτούν ότι αν θέλουν τα ρομπότ να αποτελούν κομμάτι της κοινωνίας μας, θα πρέπει να τα προγραμματίσουν με τέτοιο τρόπο, ώστε να τηρούν τις ηθικές αρχές που με αυτές ζει όλη η κοινωνία.

Είναι πια δεδομένο ότι διανύουμε την τέταρτη βιομηχανική επανάσταση. Έξυπνες μηχανές, ρομπότ που σκέφτονται, που σώζουν ανθρώπους από συντρίμια, γεωργικές μονάδες που καλλιεργούνται ψηφιακά, ρομπότ που κάνουν διάγνωση της ασθένειάς μας ακόμα και εξ αποστάσεως, αυτοκίνητα που δε χρειάζονται οδηγό, αεροπλάνα χωρίς πιλότους.

Η ανερχόμενη «τεχνολογική επανάσταση» προσφέρει μεταξύ άλλων, σπουδαίες δυνατότητες ανάπτυξης και καλύτερη ποιότητα ζωής.. Είναι αναγκαία η δημιουργία ενός αναπτυξιακού πλαισίου, με «βασικούς κανόνες και γενικές προδιαγραφές» οι οποίες θα παρέχουν ένα διαφανές, αξιόπιστο και ασφαλές περιβάλλον. Αυτό που χρειάζεται να γίνει από την πλευρά πολιτείας, είναι να χαρακτηρίσει μία στρατηγική που θα ελέγχει την αγορά και τις επιχειρήσεις καθώς και να δημιουργήσει τις απαραίτητες προϋποθέσεις ώστε να γνωρίσουν όλοι οι πολίτες τις δυνατότητες της τεχνολογίας, τις ευκαιρίες που μπορεί να προσφέρει, αλλά και να ενημερωθούν για τους πιθανούς κινδύνους που θα επιφέρει η έλευσή της.

Ο άνθρωπος μπορεί να δημιουργήσει κάτι που είναι πολύ ανώτερο από τον ίδιο. Αυτός είναι και ο λόγος που φοβάται. Φοβάται την ευφυΐα που, αν και δημιουργήθηκε από αυτόν, δεν μπορεί να την ελέγξει. Έχει ελεύθερη βούληση και γι' αυτό μπορεί να κάνει ό,τι κάνει κι αυτός. Να επαναστατήσει ενάντια στον δημιουργό του. Κι αφού είναι τόσο εύκολο ένα έξυπνο υπολογιστικό σύστημα να υπερ-πολλαπλασιαστεί αστραπιαία από μόνο του, ενώ το ανθρώπινο γένος μαστίζεται από υπογεννητικότητα, ποιος μπορεί να υποσχεθεί ότι τα ρομπότ δε θα κατακλίσουν τον κόσμο και δε θα γίνουν οι αφέντες του ανθρώπινου είδους; Τι θα τα εμποδίσει να αντιμετωπίσουν στο μέλλον τους ανθρώπους, ακόμα και σαν κατοικίδια;

Τα «συστήματα τεχνητής νοημοσύνης», είναι αναγκαίο να «εκπαιδεύονται» και να «μαθαίνουν», με σκοπό να συμβαδίζουν με τους βασικούς ηθικούς κανόνες που χαρακτηρίζουν την ανθρώπινη κοινωνία. Τα ηθικά ζητήματα που προκύπτουν, μπορούν να εκφραστούν μέσα από μία σειρά ερωτημάτων. Χαρακτηριστικό είναι το παράδειγμα με το «αυτόνομο αυτοκίνητο», το οποίο αντιλαμβάνεται ότι θα συγκρουστεί με έναν πεζό. Υποχρέωσή του είναι να πάρει γρήγορα την απόφαση εάν θα «θέσει σε κίνδυνο τους επιβάτες του» ή εάν θα «αποφύγει τον πεζό». Σε περιπτώσεις τέτοιων ακραίων καταστάσεων, όπως η παραπάνω, η απόφαση που θα πάρει αυτό το σύστημα τεχνητής νοημοσύνης, είναι αναγκαίο να είναι «προκαθορισμένη» και «γενικά αποδεκτή», καθώς και να παρέχεται η δυνατότητα ανάλυσης του ιστορικού της ενέργειας/απόφασης που έπαιρνε κάθε φορά το αυτοκίνητο.

Αξιοσημείωτο είναι ότι, η υπολογιστική νοημοσύνη συγκαταλέγεται στις πιο σημαντικές τεχνολογικές τάσεις του αιώνα. Η ραγδαία αύξηση της χρήσης και ανάπτυξής της θέτει την κοινωνία ενώπιον νέων και δύσκολων προκλήσεων. Ο δρόμος προς μια κοινωνία συνύπαρξης ανθρώπων και μηχανών, επιβάλλει μια διαφορετική προσέγγιση του τρόπου δημιουργίας των τεχνολογικών επιτευγμάτων, καθώς τα ρομπότ και η τεχνητή νοημοσύνη αναμένεται να ενισχύσουν την αλληλεπίδρασή τους με τους ανθρώπους σε πολλούς και διαφορετικούς τομείς στο μέλλον. Για να αποφευχθεί το σκοτεινό σενάριο του Έλον Μασκ, που καταλήγει με την πρόβλεψη ότι *«η υιοθέτηση της τεχνητής νοημοσύνης είναι σαν να κάνεις επίκληση στο διάβολο»*, πρέπει *«οι έρευνες για την τεχνητή νοημοσύνη να τεθούν έγκαιρα υπό έλεγχο και ρυθμιστική εποπτεία»*.

Είναι εύκολο να καταλάβει κανείς ότι ένα εξελιγμένο «υπολογιστικό σύστημα» που είναι φτιαγμένο βάση της ανθρώπινης λογικής, ύστερα από κάποιο χρονικό διάστημα, θα θελήσει να επωφεληθεί των δυνατοτήτων του και να παίρνει μόνο του κάποιες από τις αποφάσεις του. Κάποιες από αυτές τις αποφάσεις, μπορεί να οδηγήσουν στην καταστροφή της Ανθρωπότητας, αλλά κάποιες άλλες, ίσως να οδηγήσουν στη σωτηρία της. Έτσι μπορεί κανείς να αναρωτηθεί εάν μπορεί ένα ρομπότ να θεωρηθεί νομικά υπεύθυνο για τις πράξεις του. Θα πρέπει να υλοποιηθεί η αναγνώρισή του ως ένα "ηλεκτρονικό πρόσωπο" με νομική υπόσταση; Μπορούν να επιβληθούν ηθικοί κανόνες στην υπολογιστική νοημοσύνη; Με αυτά και δεκάδες ακόμα ερωτήματα, καταπιάστηκε αυτή η εργασία. Δυστυχώς, οι επιστήμονες και οι φιλόσοφοι, αν και έδωσαν τεκμηριωμένες απαντήσεις, πολύ συχνά ήταν αντικρουόμενες, υποκειμενικές και υποθετικές.

Τέλος, είναι γενικά αποδεκτό ότι η υπολογιστική νοημοσύνη μπορεί να λύσει άλυτα προβλήματα. Για την αποφυγή προβλημάτων οι στόχοι της μηχανής θα πρέπει να ευθυγραμμίζονται με τους στόχους των ανθρώπων, μέσω αξιών και δεοντολογικής συμπεριφοράς (Russel και Norvig, 2010). Από εκεί και πέρα η ανθρωπότητα θα έχει την δυνατότητα να επαναπροσδιορίσει αναλόγως τις αξίες της, να αλλάξει και να ενισχύσει την ηθική συμπεριφορά της και να επανεξετάσει την συμβολή τους στην κοινωνία σε ένα βαθύτερο βαθμό. Οι άνθρωποι πρέπει να είναι μέρος αυτής της ανάπτυξης, ώστε η ΑΙ να βοηθήσει την ενίσχυση της ανθρώπινης ευστροφίας και να δημιουργήσει μία κουλτούρα βασισμένη στην συνεργασία.

Οι αρνητικοί παράγοντες μπορούν να αποτελέσουν μία πραγματική απειλή για την εθνική ασφάλεια και μπορούν να απελευθερώσουν μία πορεία δράσης με μακροχρόνιες συνέπειες. Το ίδιο μπορεί να γίνει και με την λανθασμένη χρήση της τεχνολογίας. Η λύση βρίσκεται στην σύνθετη φύση της ηθικής. Η υπολογιστική νοημοσύνη θα πρέπει να αποτελεί ύψιστη προτεραιότητα.

Μελλοντικός στόχος είναι η περεταίρω έρευνα με περισσότερες ερευνητικές εργασίες σχετικά με την δημιουργία ηθικών ρυθμιστικών κανόνων, που θα προσδιορίζουν τις σχέσεις ανθρώπων και υπολογιστικών συστημάτων, και θα καθορίζουν πρότυπα κανονισμών ασφαλείας, την ασφάλειά τους από την ανθρωπότητα και την ασφάλεια της ανθρωπότητας από αυτά.

ΒΙΒΛΙΟΓΡΑΦΙΑ:

ΚΕΦΑΛΑΙΟ 1:

1. Vinge, V., (1993) «The Coming Technological Singularity: How to Survive in the Post-Human Era», *Department of Mathematical Sciences (San Diego State University)*. The original version of this article was presented at the VISION-21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute, March 30-31, 1993
<<https://edoras.sdsu.edu/~vinge/misc/singularity.html>>.
<<https://adsabs.harvard.edu/abs/1993vise.nasa...11V>>.

ΚΕΦΑΛΑΙΟ 2:

2. Gardner, H., (1993) «Frames of Mind: The Theory of Multiple Intelligences», New York: Basic Books.
3. Niven, L., (1967) «The Ethics of Madness», σελ:82-108.
Reprinted in *Neutron Star*, Larry Niven, Ballantine Books, 1968.
4. Ταξίδι στην αρχαία Ελλάδα: Οι Σημαντικότερες Εφευρέσεις Των Αρχαίων Ελλήνων,
<http://archaia-ellada.blogspot.com/2014/03/blog-post_2378.html>.
5. Rasmussen, S. et al., (1991) «Computational Connectionism within Neurons: a Model of Cytoskeletal Automata Subserving Neural Networks, in *Emergent Computation*, Stephanie Forrest», σελ: 428-449, MIT Press.
6. Conrad, M., et al., (1989) «Towards an Artificial Brain, *BioSystems*», vol. 23, σελ:175-218.
7. Sims, Karl, (1991) «Interactive Evolution of Dynamical Systems, *Thinking Machines Corporation, Technical Report Series*» Published in *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*, Paris: MIT Press.
8. Hass, A., (1998) «Doing the eight thing: Cultivating your moral intelligence».
9. Kline, Ronald R., *Cybernetics*, (2011) «Automata Studies and the Dartmouth Conference on Artificial Intelligence», *IEEE Computer Society*
<https://www.computer.org/web/about>
10. McCarthy, J., Minsky, M., Rochester, N., Shannon, C.E., (1955) «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence».,
<<http://jmc.stanford.edu/articles/dartmouth.html>>.
11. Covey, S.R., (1989) «The 7 Habits of Highly Effective People». New York: Fireside Book.
12. Διάλεξη ICSD :: 2008 - 2018 Πανεπιστήμιο Αιγαίου, Υπολογιστική Νοημοσύνη
13. Λυκοθανάσης, Σ., (2001) «Γενετικοί Αλγόριθμοι και Εφαρμογές», εκδόσεις ΕΑΠ.
14. Αργυράκης, Π., (2001) «Νευρωνικά Δίκτυα και Εφαρμογές», εκδόσεις ΕΑΠ.
15. Aleksander, I. & Morton, H., (1990) «An Introduction to Neural Computing», ISBN:0-442-31218-0
<<https://dl.acm.org/citation.cfm?id=131391>>.
16. Genetic Algorithms in Search, Optimization, and Machine Learning
<https://www2.fiit.stuba.sk/~kvasnicka/Free%20books/Goldberg_Genetic_Algorithms_in_Search.pdf>.
17. Bezdek, J., (1994) «What is computational intelligence? In: *Computational Intelligence*»,
<<https://www.osti.gov/biblio/81587>>.
18. http://www.icsd.aegean.gr/lecturers/kavallieratou/NN&EP_files/ci_1.pdf

19. De Jong, (2006) «Evolutionary Computation: A Unified Approach», *MIT Press*.
<https://epdf.tips/evolutionary-computation-a-unified-approach.html>
20. Kuhn, A., (2016) «Artificial Intelligence, Computational Intelligence, SoftComputing, Natural Computation - what's the difference?»,
<<https://www.andata.at/en/answer/artificial-intelligence-computational-intelligence-softcomputing-natural-computation-whats-the-difference.html>>.

ΚΕΦΑΛΑΙΟ 3:

21. Bezdek, J., (1992) «On the relationship between Neural Networks, Pattern Recognition and Intelligence Division of Computer Science», *University of West Florida, Pensacola*,
<<https://core.ac.uk/download/pdf/82489747.pdf>>.
22. Χαδιαράκου, Ε., Borba, M., (2010) «Building Moral Intelligence (Χτίζοντας την Ηθική Νοημοσύνη)»,
<<https://www.in.gr/2009/02/11/life/kid/ithiki-noimosyni-to-epiktito-xarisma/>>.
23. Computer ethics definitions, <<https://techterms.com/definition/computerethics>>.
24. Bostrom, N., Yudkowsky, Eliezer, (2014) «The ethics of artificial intelligence». Επιμ. The Cambridge Handbook of Artificial Intelligence. Cambridge: Cambridge University
<<https://www.cambridge.org/core/books/cambridge-handbook-of-artificial-intelligence/ethics-of-artificial-intelligence/B46D2A9DF7CF3A9D92601D9A8ADA58A8>>.
25. Easton, (2015) «Roboethics – We need Universal Robots Rights, Ethics and Legislation»,
<<https://wtvox.com/robotics/roboethics-universal-robot-rights-ethics-and-legislation/>>.
26. Δεληγιάννης, Κ., (2018) «Τεχνητή νοημοσύνη: Ευλογία ή κατάρα;»,
<<http://www.kathimerini.gr/795998/article/tehnologia/computers/tehnth-noimosynh-eylogia-h-katara>>.
27. Asaro, P., (2006) «What should we want from a robot ethic»,
– all rights reserved www.i-r-i-e.net, *International Review of Information Ethics*,
<<http://peterasaro.org/writing/Asaro%20IRIE.pdf>>.
28. Sütfeld, L., Gast, R., König, P., και Pipa, G., (2017) «Using Virtual Reality to Assess Ethical Decisions in Road Traffic Scenarios: Applicability of Value-of-Life-Based Models and Influences of Time Pressure», *Institute of Cognitive Science*,
<journal.frontiersin.org/article/10.3389/fnbeh.2017.00122/full>.
29. Floridi, L., (1998) «Information ethics: On the philosophical foundation of computer ethics»
USA: Publisher Kluwer Academic Publishers Hingham
<<https://dl.acm.org/citation.cfm?id=594590>>.
30. Newsroom, CNN Greece, (2016) «Έτοιμη να σκέφτεται και να εξελίσσεται η τεχνητή νοημοσύνη»,
<<https://www.cnn.gr/tech/story/22844/etoimi-na-skeftetai-kai-na-exelissetai-i-texniti-noimosyni>>.
31. Newsbomb.gr, (2015) «Τρόμος: ρομπότ με πολύ υψηλή νοημοσύνη δραπέτευσε για δεύτερη φορά από εργαστήριο»,
<<https://www.newsbomb.gr/kosmos/story/707230/tromos-rompot-me-poly-ypsili-noimosyni-drapeteuse-gia-deyteri-fora-apo-ergastirio-kai-video#ixzz4iYEumzie>>.
32. Naftemporiki.gr, (2015) «Στίβεν Χόκινγκ, έλον Μασκ και άλλοι 1000 ειδικοί κατά της στρατιωτικής τεχνητής νοημοσύνης»,
<<https://www.naftemporiki.gr/story/983065/stiben-xokingk-elon-mask-kai-alloi-1000-eidikoi-kata-tis-stratiotikis-tenitis-noimosunis>>.
33. Honam, D., (2016) «A Robot's Bill of Rights»,
<<https://bigthink.com/think-tank/a-robots-bill-of-rights>>.
34. Wolchover, N., (2015) «Concerns of an Artificial Intelligence Pioneer»,

<<https://www.quantamagazine.org/artificial-intelligence-aligned-with-human-values-qa-with-stuart-russell-20150421/>>.

35. McDonald, C., (2015) «The Good, The Bad and The Robot: Experts Are Trying to Make Machines Be “Moral” », <<https://alumni.berkeley.edu/california-magazine>>.

ΚΕΦΑΛΑΙΟ 4:

36. Mavridis, N., (2011) «A thousand eyes, a thousand hands», TEDx Athens <<https://www.youtube.com/watch?v=81bWAwVUj5Q>>.
37. IEEE International Conference, (2011)
38. Μαυρίδης, Ν., (2015) «Ρομποτική, Τεχνητή Νοημοσύνη, Αισθητική και Ηθική», <https://www.onassis.org/onassis-magazine/issue-69/the_beauty_of_ethical_robots>.
39. ΑΠΕ-ΜΠΕ, (2018) «Η συγκλονιστική διάλεξη του 37χρονου καθηγητή – ιδιοφυία Κωνσταντίνου Δασκαλάκη στο ΑΠΘ», <<https://www.iefimerida.gr/news/389729/i-sygklonistiki-dialexi-toy-37hronoy-kathigiti-idiofyia-konstantinoy-daskalaki-sto-apth#ixzz56v3L4Mjy>>.
40. Markoff, J., (2009) «Scientists Worry Machines May Outsmart Man», <https://www.nytimes.com/2009/07/26/science/26robot.html?_r=1&module=ArrowsNav&contentCollection=Science&action=keypress®ion=FixedLeft&pgtype=article>.
41. ΑΠΕ-ΜΠΕ, (2017) «Αλγόριθμοι τεχνητής νοημοσύνης ξεχωρίζουν τους "γκέι" από τους "στρέιτ"», <<https://www.elftheriaonline.gr/ellada-kosmos/tech/item/133373-algorithmoi-texnitis-noimosynis-ksexorizoun-tous-gkei-apo-tous-streit>>.
42. Vincent, J., (2016) «Mall security bot knocks down toddler, breaks Asimov’s first law of robotics», <<https://www.theverge.com/2016/7/13/12170640/mall-security-robot-k5-knocks-down-toddler>>.
43. Το Φανάρι, (2018) «Τεχνητή νοημοσύνη διαβάζει το μυαλό σας και αναπαράγει τις εικόνες», <https://tofani.blogspot.com/2018/01/blog-post_42.html>.
44. Καψύλης, Α., (2016) «Η τεχνητή νοημοσύνη σύμμαχος και ανταγωνιστής του ανθρώπου», <<https://www.tovima.gr/>>.
45. Οικονόμου, Τ., (2016) «Ασφαλιστικές εισφορές για τα ρομπότ προβλέπει Ευρωπαϊκή πρόταση», <<https://www.cnn.gr/tech/story/36556/asfalistikes-eisfores-gia-ta-rompot-provlepei-eyropaiki-protasi>>.
46. Μπουκουβάλα, Χ., (2017) «Οι μάσκες μίας απολυταρχικής εξουσίας και η μοναδικότητα», <<https://aperopia.fr/el/>>.
47. Ασκητής, Θ., Μόσχα, Μ., (2017) «Οι γυναίκες θα κάνουν περισσότερο σεξ με ρομπότ παρά με άνδρες το 2025!», <www.healthnews.com.cy>.
48. ΑΠΕ-ΜΠΕ, (2017) «Ρομποτικό χέρι», <<https://gr.euronews.com/2017/05/04/bionic-hand-with-artificial-ntelligencei>>.
49. Kuhn, R., L., (2015) «When Robots Colonize the Cosmos, Will They Be Conscious? (Op-Ed)», <https://www.space.com/30937-when-robots-colonize-cosmos-will-they-be-conscious.html?cmpid=514630_20151107_54648356&adbid=10153144907361466&adbpl=f>

[b&adbpr=17610706465&fbclid=IwAR1GR0sgjXle0y7aawErQtUsq-BSFt-htfLQG4KjloI6ev8kln15E5IBMy4>.](https://www.nickbostrom.com/superintelligence.html)

50. Bostrom, N., (2006) «How long before Superintelligence?», *Linguistic and Philosophical Investigations*, 5(1), σελ:11-30.
<<https://www.nickbostrom.com/superintelligence.html>>.
51. Bostrom, N., (2002) «Existential risks: Analyzing Human Extinction Scenarios and Related Hazards», *Journal of Evolution and Technology*, 9(1)
<<https://www.nickbostrom.com/existential/risks.html>>.
52. Freitas Jr., R. A. Nanomedicine, (1999), Volume 1: «Basic Capabilities». Georgetown: Landes Bioscience.
<<https://www.nanomedicine.com>>.
53. Hanson, R., et al. (1998), «A Critical Discussion of Vinge's Singularity Concept», Extropy Online.
<<https://www.extropy.org/eo/articles/vi.html>>.

ΚΕΦΑΛΑΙΟ 5

54. Moravec, H., (1999) «Robot: Mere Machine to Transcendent Mind». New York: Oxford University.
55. Yudkowsky, E., (2002) «The AI Box Experiment»,
<<https://sysopmind.com/essays/aibox.html>>.
56. Yudkowsky, E., (2003) «Creating Friendly AI»,
<<https://www.singinst.org/CFAI/index.html>>.
57. Chalmers, D., (2010) «Singularity, Intelligence Explosion»,
<<https://singinst.org/blog/2010/04/08/david-chalmers-on-singularity-intelligence-explosion/>>.
58. Loosemore, R., και Goertzel, B., (2011) «Why an Intelligence Explosion is Probable»,
<<https://hplusmagazine.com/2011/03/07/why-an-intelligence-explosion-is-probable/>>.
59. OHeigeartaigh, S., (2013) «Would you hand over a moral decision to a machine? Why not? Moral outsourcing and Artificial Intelligence»,
<<https://blog.practicaethics.ox.ac.uk/2013/08/would-you-hand-over-a-moral-decision-to-a-machine-why-not-moral-outsourcing-and-artificial-intelligence/>>.
60. Μαυρίδης, Ν., (2014) «Τι χρειάζεται για να δοθεί σε κάποιον η άδεια φόνου;»,
<https://www.dr-nikolaos-mavridis.com/resources/RoboCop_Mavridis.pdf>.
61. Damasi, A., (2000) «Το λάθος του Καρτέσιου. Συγκίνηση, Λογική και Ανθρώπινος Εγκέφαλος.», Μετάφραση Κώστας Παπακωνσταντίνου, σελ: 144,
<<https://www.protoporia.gr/>>.
62. Χαράλαμπος, Κ., (2017) «Πρόταση σοκ για τον πλανήτη – ποιον θέλουν να φέρουν για να κυβερνήσει»,
<<https://katohika.gr/>>.
63. Ρέμπελου, Α., (2015) «Google: επικίνδυνα παιχνίδια με ρομπότ»,
<<https://www.news247.gr/weekend-edition/google-epikindyna-paichnidia-me-rompot.6362919.html>>.
64. Newsroom, CNN Greece, (2017) «Σοφία το ρομπότ που πήρε υπηκοότητα από την Σαουδική Αραβία»,
<<https://www.cnn.gr/tech/story/103290/h-sofia-to-rompot-pire-ypikootita-apo-ti-saoydiki-aravia-vid>>.

65. ΑΠΕ-ΜΠΕ, (2018) «Η συγκλονιστική διάλεξη του 37χρονου καθηγητή ιδιοφυΐα Κωνσταντίνου Δασκαλάκη στο ΑΠΘ»,
<<https://www.iefimerida.gr/news/389729/i-sygklonistiki-dialexi-toy-37hronoy-kathigiti-idiofyia-konstantinoy-daskalaki-sto-apth#ixzz56v3L4Mjy%2016|01|2018%2023:12>>.
66. Ομιλία του KURZWEIL, R., (2011) στο TED «Ένα πανεπιστήμιο για την επερχόμενη μοναδικότητα»,
<<https://antikleidi.com/2011/11/04/2045-immortal/>>.
67. White, J., (2016) «Ethics of Artificial Intelligence», Okinawa Institute of Science And Technology,
<<https://philpapers.org/browse/ethics-of-artificial-intelligence/>>.

ΚΕΦΑΛΑΙΟ 6:

68. Μαυρίδης, Ν., (2015) «Ρομποτική, Τεχνητή Νοημοσύνη, Αισθητική και Ηθική», τεύχος 69,
<https://www.onassis.org/onassis-magazine/issue-69/the_beauty_of_ethical_robots>.
69. Μαυρίδης, Ν., (2014) «Τι χρειάζεται για να δοθεί σε κάποιον η άδεια φόνου;»,
<<https://greeklish.info/gr/world/262>>.
70. Κρασαδάκης, Γ., (2018) «Τεχνητή Νοημοσύνη: Οι κίνδυνοι»,
<<https://www.naftemporiki.gr/story/1314766/texniti-noimosuni-oi-kindunoi>>.
71. HUMAN Right Council, (2013) «U.N. Report Calling for Moratoria on Lethal Robots»,
<<https://news.usni.org/2013/05/03/document-u-n-report-calling-for-moratoria-on-lethal-robots>>.
72. The Economics Times, (2013) «India developing robotic soldiers to replace humans in warfare»,
<<https://economictimes.indiatimes.com/india-developing-robotic-soldiers-to-replace-humans-in-warfare/articleshow/20553020.cms>>.
73. (2017) National Geographic Documentary: Year Million (Artificial Intelligence),
<https://www.youtube.com/watch?v=Xn6IDytVGw&fbclid=IwAR1FLAIt1UJ4J6b_ODsjR4XntiNRlKU5Wf54fGvNX74I37rFDInfA72WCC0>.
74. Anderson, K., and Matthew C. Waxman, (2012) «Law and Ethics for Robot Soldiers»,
<<https://www.cfr.org/node/161025>>.
75. Bridle, J., (2018) «Rise of the machines: has technology evolved beyond our control?»
<https://www.theguardian.com/books/2018/jun/15/rise-of-the-machines-has-technology-evolved-beyond-our-control-?utm_source=esp&utm_medium=Email&utm_campaign=Bookmarks%20-%20Collections%202017&utm_term=278296&subid=2661253&CMP=bookmarks_collection&fbclid=IwAR2MctkeYpHWpF1VnCX3SYOaqE8tiG726ctjg8kj5Smf3h0s89u4ocS77A>.
76. Ricks, T., (2013) «Rosa’s Dystopia: The Moral Downside of Coming Autonomous Weapons Systems»,
<<https://foreignpolicy.com/2013/06/18/rosas-dystopia-the-moral-downside-of-coming-autonomous-weapons-systems/>>.
77. Carpenter, C., (2013) «Προσοχή στα ρομπότ δολοφόνους»,
<<https://foreignaffairs.gr/articles/69393/charli-carpenter/prosoxi-sta-rompot-dolofonoys?page=show>>.
78. Suarez, D., TEDGlobal (2013) «The kill decision shouldn’t belong to a robot»,

https://www.ted.com/talks/daniel_suarez_the_kill_decision_shouldn_t_belong_to_a_robot?fbclid=IwAR1YGbvHHClfTP95LUfgc95wQtQowDbabogDaywjA4lseLSot39qKSoBuOO

ΚΕΦΑΛΑΙΟ 7:

79. Scott, A., (2015) «Human Rights for Robots»,
<<https://blog.dilbert.com/post/116298066321/human-rights-for-robots>>.
80. Robertson, J., (2014) «Human rights vs robot rights»,
<<https://criticalasianstudies.org/issues/vol46/no4/human-rights-vs-robot-rights.html>>.
81. Καθημερινή, (2017) «Φορολόγηση των ρομπότ προτείνει ο Μπιλ Γκέιτς»,
<<https://www.kathimerini.gr/897214/article/epikairothta/kosmos/forologhsh-twn-rompot-proteinei-o-mpil-gkeits>>.
82. Millner, J., (2015) «Should robots have human rights? Act now to regulate killer machines before they multiply and demand the right to vote, warns legal expert»,
<<https://www.dailymail.co.uk/sciencetech/article-3168081/Should-robots-human-rights-Act-regulate-killer-machines-multiply-demand-right-vote-warns-legal-expert.html>>.
83. Delvaux, M., (2016) (Πρωτοβουλία- άρθρο 36 του Κανονισμού), Ευρωπαϊκό Κοινοβούλιο 2014-2019, Επιτροπή Νομικών Θεμάτων, «Σχέδιο Έκθεσης που περιέχει συστάσεις προς την Επιτροπή για ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής »,2015/2013(INL).
84. Αναστάσιος, Α., (2016) «Lex Machina»,
<https://www.huffingtonpost.gr/anastasios-antoniou/lex-machina_b_8624316.html>.
85. ΣΚΑΪ, (2017) «Τεχνητή νοημοσύνη: Η ΕΕ θα αναγνωρίζει την προσωπικότητα των ρομπότ»,
<<http://www.skai.gr/news/technology/article/336348/tehni-noimosuni-i-ee-tha-anagorizeitin-prosopikotita-ton-robot/>>.
86. Lawspot.gr, (2018) «Τεχνητή νοημοσύνη και ρομποτική: Ηθικές και νομικές προκλήσεις»,
<<https://www.lawspot.gr/nomika-nea/tehni-noimosyni-kai-rompotiki-ithikes-kai-nomikes-prokliseis>>.
87. (2017), Έκθεση με συστάσεις προς την Επιτροπή σχετικά με ρυθμίσεις αστικού δικαίου στον τομέα της ρομποτικής 2015/2103(INL)),
<<http://www.europarl.europa.eu>>.
88. Ashrafian, H., (2015) Science and Engineering Ethics «AIonAI: A Humanitarian Law of Artificial Intelligence and Robotics»,
<https://www.researchgate.net/publication/259700270_AIonAI_A_Humanitarian_Law_of_Artificial_Intelligence_and_Robotics>.
89. BBC NEWS, (2006) «Robots could demand legal rights»,
<<http://news.bbc.co.uk/2/hi/technology/6200005.stm>>.
90. Ashrafian, H., (2015) «Intelligent robots must uphold human rights», Nature 519, 391 doi:10.1038/519391a.
<<https://www.nature.com/news/intelligent-robots-must-uphold-human-rights-1.17167>>.
91. <<http://www.zimzamphysics.gr/>>
92. <<https://www.huffingtonpost.gr/> , <https://www.huffingtonpost.gr/>>